

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 08:18:56 ; Search time 10 seconds  
(without alignments)  
240.479 Million cell updates/sec

Title: US-09-701-623C-5  
Perfect score: 140  
Sequence: 1 CGETQSRVTHPLPALMRSTTKC 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	117	83.6	428	1 EHHU	Ig epsilon chain C
2	113	80.7	426	2 I36948	Ig epsilon-chain -
3	80	57.1	429	1 EHRT	Ig epsilon chain C
4	56	40.0	107	2 I68730	IgE chain C3 regio
5	56	40.0	107	2 I68726	IgE chain C3 regio
6	56	40.0	388	1 EHMS	Ig epsilon chain C
7	56	40.0	548	2 S38864	Ig epsilon chain C
8	54	38.6	684	2 S60266	novel antigen rece
9	52	37.1	320	2 C86148	hypothetical prote
10	50	35.7	1597	2 S65053	genome polyprotein
11	50	35.7	1601	2 S48699	178K protein - tob
12	49.5	35.4	1235	1 S16948	insulin receptor s
13	48	34.3	388	2 B31194	hypothetical prote
14	48	34.3	1065	2 T52054	cellulose synthase
15	48	34.3	2254	2 D86215	protein T6D22.14 [
16	47.5	33.9	381	2 S28115	gas-vesicle protei
17	47.5	33.9	382	2 JQ1122	gas-vesicle protei
18	47.5	33.9	382	2 T08243	gas-vesicle operon
19	47	33.6	242	1 MEIVCJ	matrix protein M1
20	47	33.6	1940	1 S04090	myosin heavy chain
21	47	33.6	1940	1 A24322	myosin heavy chain
22	47	33.6	1940	2 S20320	myosin heavy chain
23	46.5	33.2	1231	2 S30185	insulin receptor s
24	46	32.9	116	2 S37909	hypothetical prote
25	46	32.9	247	2 A27547	trypsin (EC 3.4.21
26	46	32.9	430	2 AG0531	cell cycle protein
27	46	32.9	550	1 VGBEL8	glycoprotein E - h
28	46	32.9	842	2 C83458	conserved hypothet
29	46	32.9	1019	2 A83613	conserved hypothet

## ALIGNMENTS

### RESULT 1

EHHU

Ig epsilon chain C region - human

C/Species: Homo sapiens (man)

C/Date: 31-Mar-1981 #sequence revision 13-Jun-1983 #text change 16-Jul-1999

C/Accession: A22771; A23195; PH1214; A93491; A90824; A94418; B93933; S02438; A53116; C41

R/Flanagan, J.G.; Rabbitts, T.H.

EMBO J. 1, 655-660, 1982

A/Title: The sequence of a human immunoglobulin epsilon chain constant region gen

A/Reference number: A22771; MUID:84236029; PMID:6234164

A/Accession: A22771

A/Molecule type: DNA

A/Residues: 1-428 <FLA>

A/Cross-references: GB:L00022; GB:J00227; GB:V00555; NID:g185035

R/Ueda, S.; Nakai, S.; Nishida, Y.; Hisajima, H.; Honjo, T.

EMBO J. 1, 1533-1544, 1982

A/Title: Long terminal repeat-like elements flank a human immunoglobulin epsilon pseudo

A/Reference number: A23195; MUID:84207910; PMID:6327276

A/Accession: A23195

A/Molecule type: DNA

A/Residues: 2-428 <UED>

A/Cross-references: GB:J00222; NID:g184755

R/Zhang, K.; Saxon, A.; Max, E.E.

J. Exp. Med. 176, 233-243, 1992

A/Title: Two unusual forms of human immunoglobulin E encoded by alternative RNA splicin

A/Reference number: PH1214; MUID:92308839; PMID:1613458

A/Accession: PH1214

A/Molecule type: DNA

A/Residues: 320-428 <ZHA>

A/Cross-references: EMBL:X63693; GB:S38668; NID:g32987

R/Seno, M.; Kurokawa, T.; Ono, Y.; Onda, H.; Sasaki, R.; Igarashi, K.; Kikuchi, M.; Sug

Nucleic Acids Res. 11, 719-726, 1983

A/Title: Molecular cloning and nucleotide sequencing of human immunoglobulin epsilon ch

A/Reference number: A93491; MUID:83168897; PMID:6300763

A/Accession: A93491

A/Molecule type: mRNA

A/Residues: 1-428 <SEN>

A/Cross-references: GB:L00022; GB:J00227; GB:V00555; NID:g185035

R/Max, E.E.; Battey, J.; Ney, R.; Kirsch, I.R.; Leder, P.

Cell 29, 691-699, 1982

A/Title: Duplication and deletion in the human immunoglobulin epsilon genes.

A/Reference number: A90824; MUID:83001945; PMID:6288268

A/Accession: A90824

A/Molecule type: DNA

A/Residues: 1-358, 'L', 360-428 <MAX>

A/Cross-references: GB:J00222; NID:g184755

A/Note: this sequence difference may be due to polymorphism

R/Bennich, H.H.; Johansson, S.G.O.; Von Bahr-Lindstrom, H.

in Immediate Hypersensitivity: Modern Concepts and Developments, Bach, M.K., ed., pp.1-

A/Reference number: A94418

A/Accession: A94418

A/Molecule type: protein

A;Residues: 'GAWTL', 6, 'X', 8-16, 'B', 18-43, 'B', 45-52, 55-92, 95-97, 'B', 99-121, 'B', 123, 'L', 12  
A;Experimental source: myeloma protein Nd  
R;Kenter, J.H.; Molgaard, H.V.; Houghton, M.; Derbyshire, R.B.; Viney, J.; Bell, L.O.; G  
Proc. Natl. Acad. Sci. U.S.A. 79, 6661-6665, 1982  
A;Title: Cloning and sequence determination of the gene for the human immunoglobulin epsilon  
A;Reference number: A93933; MUID: 83065234; PMID: 6815656  
A;Accession: E93933  
A;Molecule type: mRNA  
A;Residues: 1-40; 68-114; 427-428 <XEN>  
A;Cross-references: GB:L00022; NID:g185035  
R;Ikayama, S.  
FEBS Lett. 224, 306-310, 1987  
A;Title: Purification and characterization of a recombinant human IgE Fc-epsilon fragment  
A;Reference number: S02438; MUID: 88093554; PMID: 3121397  
A;Accession: S02438  
A;Status: nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 98-352 <IKE>  
R;Zhang, K.; Max, E.E.; Chesh, H.K.; Saxon, A.  
J. Biol. Chem. 269, 456-462, 1994  
A;Title: Complex alternative RNA splicing of epsilon-immunoglobulin transcripts produces  
A;Reference number: A53116; MUID: 94103254; PMID: 8276835  
A;Accession: A53116  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 320-428 <ZH2>  
A;Experimental source: myeloma U266-derived cell line AF-10  
A;Note: sequence extracted from NCBI backbone (NCBIN:141701, NCBIP:141702)  
R;Hellman, L.  
Eur. J. Immunol. 23, 159-167, 1993  
A;Title: Characterization of four novel epsilon chain mRNA and a comparative analysis of  
A;Reference number: A46536; MUID: 93122085; PMID: 8419166  
A;Accession: A46536  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 382-426 <HEL>  
A;Cross-references: GB:S55273; NID:g263166; PIDN:AAB24857.1; PID:g263167  
A;Experimental source: B cell myeloma U-266  
A;Note: sequence extracted from NCBI backbone (NCBIP:125297)  
A;Accession: A46536  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 382-391 <HE2>  
A;Cross-references: GB:S55276; NID:g263168; PIDN:AAB24858.1; PID:g263169  
A;Experimental source: B cell myeloma U-266  
A;Note: sequence extracted from NCBI backbone (NCBIP:125299)  
A;Accession: A46536  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: DNA  
A;Residues: 401-428 <HE3>  
A;Cross-references: GB:S53497; NID:g263162; PIDN:AAB24855.1; PID:g263163  
A;Experimental source: B cell myeloma U-266  
A;Note: sequence extracted from NCBI backbone (NCBIP:123483)  
C;Genetics:  
A;Gene: GDB:IGHE  
A;Cross-references: GDB:119335; OMIM:147180  
A;Map position: 14q32.33-14q32.33  
A;Introns: 1/1; 104/1; 211/1; 319/1  
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap  
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la  
C;Superfamily: immunoglobulin C region; immunoglobulin homology  
C;Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglob  
F;22-87/Domain: immunoglobulin homology <IM1>  
F;128-195/Domain: immunoglobulin homology <IM2>  
F;232-301/Domain: immunoglobulin homology <IM3>  
F;338-407/Domain: immunoglobulin homology <IM4>  
F;14/Disulfide bonds: interchain (co light chain) #status predicted  
F;15-105, 29-85, 135-193, 239-299, 345-405/Disulfide bonds: #status predicted  
F;21, 49, 99, 146, 252, 275/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F;121, 209/Disulfide bonds: interchain (to heavy chain) #status predicted

Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 GETYQSRVTHPHLPALMRSTTK 24  
|||||  
Db 294 GETYQCRVTHPHLPALMRSTTK 316  
|||||  
RESULT 2  
136948  
Ig epsilon-chain - chimpanzee (fragment)  
C;Species: Pan troglodytes (chimpanzee)  
C;Date: 04-Oct-1996 #sequence\_revision 04-Oct-1996 #text\_change 21-Jan-2000  
C;Accession: I36948  
R;Sakoyama, Y.; Hong, K.  
Proc. Natl. Acad. Sci. U.S.A. 84, 1080-1084, 1987  
A;Title: Nucleotide sequences of immunoglobulin epsilon genes of chimpanzee and orangut  
A;Reference number: I36948; MUID: 87147196; PMID: 3103123  
A;Accession: I36948  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-426 <RES>  
A;Cross-references: GB:M15398; NID:g176797; PIDN:AAA35416.1; PID:g176798  
C;Genetics:  
A;Introns: 103/1; 209/1; 317/1  
C;Superfamily: immunoglobulin C region; immunoglobulin homology  
F;336-405/Domain: immunoglobulin homology <IMM>  
Query Match 80.7%; Score 113; DB 2; Length 426;  
Best Local Similarity 91.3%; Pred. No. 2.4e-09;  
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 2 GETYQSRVTHPHLPALMRSTTK 24  
|||||  
Db 292 GETYQCRVTHPHLPALVRSTTK 314  
|||||  
RESULT 3  
EHRT  
Ig epsilon chain C region - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 17-Dec-1982 #sequence\_revision 17-Dec-1982 #text\_change 16-Jul-1999  
C;Accession: A93442; A90937; A02143  
R;Hellman, L.; Pettersson, U.; Engstrom, A.; Karlsson, T.; Bennich, H.  
Nucleic Acids Res. 10, 6041-6049, 1982  
A;Title: Structure and evolution of the heavy chain from rat immunoglobulin E.  
A;Reference number: A93442; MUID: 83064537; PMID: 6292865  
A;Accession: A93442  
A;Molecule type: mRNA  
A;Residues: 1-429 <HEL>  
A;Experimental source: strain LOU/c/Wsl, immunocytooma IR2  
R;Kinsvogel, W.R.; Reddy, E.P.; Moore, J.M.; Faust Jr., C.H.  
DNA 1, 335-343, 1982  
A;Title: A cloned cDNA probe for rat immunoglobulin epsilon heavy chain: construction,  
A;Reference number: A90937; MUID: 83182019; PMID: 6820340  
A;Contents: myeloma IR162  
A;Accession: A90937  
A;Molecule type: mRNA  
A;Residues: 'N', 169-307, 'L', 309-342 <KIN>  
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (ka  
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 1  
C;Superfamily: immunoglobulin C region; immunoglobulin homology  
C;Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
F;118-186/Domain: immunoglobulin homology <IM1>  
F;223-291/Domain: immunoglobulin homology <IM2>  
F;327-398/Domain: immunoglobulin homology <IM3>  
F;46, 99, 170, 240, 265, 369, 419/Binding site: carbohydrate (Asn) (covalent) #status predict  
Query Match 57.1%; Score 80; DB 1; Length 429;  
Best Local Similarity 60.9%; Pred. No. 0.00025;  
Matches 14; Conservative 3; Mismatches 6; Indels 0; Gaps 0;  
QY 2 GETYQSRVTHPHLPALMRSTTK 24

RESULT 9  
C86148  
hypothetical protein T1NC.10 - *Arabidopsis thaliana*

C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 02-Jun-2003  
 C;Accession: C86148  
 R;Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Holt, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, R.; Rizzio, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, R.; Rizzio, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, B.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A;Title: Sequence and analysis of chromosomes 1 of the plant Arabidopsis.  
 A;Reference number: A86141; MUID:21016719; PMID:11130712  
 A;Accession: C86148  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-320 <SFO>  
 A;Cross-references: GB:AE005172; NID:g8671838; PIDN:AAF78401.1; GSPDB:GN00141  
 C;Genetics:  
 A;Map position: 1  
 C;Superfamily: acyl-CoA thioesterase II (TesB)

Query Match 37.1%; Score 52; DB 2; Length 320;  
 Best Local Similarity 60.0%; Pred. No. 3.3;  
 Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 3 ETYQSRVTHPLRPA 17  
 Db 142 ELRESRITDPLRPS 156

RESULT 10  
 S65053  
 Genome polyprotein - Chinese rape mosaic virus  
 N;Alternate names: RNA replicase  
 N;Contains: RNA-directed RNA polymerase (EC 2.7.7.48)  
 C;Species: Chinese rape mosaic virus  
 C;Date: 19-Mar-1997 #sequence\_revision 29-Aug-1997 #text\_change 17-Mar-1999  
 C;Accession: S65053  
 R;Aguiar, I.; Sanchez, F.; Martin Martin, A.; Martinez-Herrera, D.; Ponz, F.  
 Plant Mol. Biol. 30, 191-197, 1996  
 A;Title: Nucleotide sequence of Chinese rape mosaic virus (oilseed rape mosaic virus), a  
 A;Reference number: S65053; MUID:96197410; PMID:8616237  
 A;Accession: S65053  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: genomic RNA  
 A;Residues: 1-1597 <AGU>  
 A;Cross-references: EMBL:U30944  
 A;Note: readthrough of the terminator UGA occurs between codons CAA for 1103-Gln and CAA  
 A;Note: the internal stop codon is translated as X  
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1995  
 C;Superfamily: cucumber mosaic virus RNA 1 protein  
 C;Keywords: nucleotidyltransferase

Query Match 35.7%; Score 50; DB 2; Length 1597;  
 Best Local Similarity 36.8%; Pred. No. 35;  
 Matches 14; Conservative 4; Mismatches 6; Indels 14; Gaps 2;

Qy 2 GETYQS-----RVT-----HPHLPRALMRSTTKC 25  
 Db 1031 GETYEKTAIVLTATPLEIIISRASPHVLVALTRHTTRC 1068

RESULT 11  
 S48699  
 178K protein - tobacco mosaic virus (strain cr-TMV)  
 N;Alternate names: readthrough protein  
 N;Contains: 122K protein  
 C;Species: tobacco mosaic virus, TMV  
 A;Variety: strain cr-TMV  
 C;Date: 06-Dec-1996 #sequence\_revision 06-Dec-1996 #text\_change 20-Sep-1999  
 C;Accession: S48699; S48659

R;Dorokhov, Y.L.; Ivanov, P.A.; Novikov, V.K.; Agranovsky, A.A.; Morozov, S.Y.; Efimov, F.B.S. Lett. 350, S-8, 1994  
 A;Title: Complete nucleotide sequence and genome organization of a tobamovirus infecting  
 A;Reference number: S48659; MUID:94341372; PMID:7545946  
 A;Accession: S48699  
 A;Molecule type: genomic RNA  
 A;Residues: 1-1601 <DOR>  
 A;Cross-references: EMBL:Z29370; NID:g488713; PIDN:CAA82559.1; PID:g615908  
 A;Experimental source: tobamovirus infecting cruciferae plants (cr-TMV)  
 A;Note: readthrough of the terminator UGA occurs between codons CAA for 1107-Gln and CAA  
 A;Note: the internal stop codon is translated as X  
 C;Superfamily: cucumber mosaic virus RNA 1 protein  
 F;1-1601/Product: 178K protein #status predicted <PRO2>  
 F;1-1107/Product: 122K protein #status predicted <PRO1>

Query Match 35.7%; Score 50; DB 2; Length 1601;  
 Best Local Similarity 36.8%; Pred. No. 35;  
 Matches 14; Conservative 4; Mismatches 6; Indels 14; Gaps 2;

Qy 2 GETYQS-----RVT-----HPHLPRALMRSTTKC 25  
 Db 1035 GETYEKTAIVLTATPLEIIISRASPHVLVALTRHTTRC 1072

RESULT 12  
 S16948  
 Insulin receptor substrate IRS-1 - rat  
 N;Alternate names: insulin receptor substrate pp185  
 C;Species: Rattus norvegicus (Norway rat)  
 C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 18-Jun-1999  
 C;Accession: S16948; A39811  
 R;Sun, X.J.; Rothenberg, P.; Kahn, C.R.; Backer, J.M.; Araki, E.; Wilden, P.A.; Cahill, Nature 352, 73-77, 1991  
 A;Title: Structure of the insulin receptor substrate IRS-1 defines a unique signal tran  
 A;Reference number: S16948; MUID:91287824; PMID:1648180  
 A;Accession: S16948  
 A;Molecule type: mRNA  
 A;Residues: 1-1235 <SUN>  
 A;Cross-references: EMBL:X58375; NID:g56503; PIDN:CAA41264.1; PID:g56504  
 R;Rothenberg, P.L.; Lane, W.S.; Karasik, A.; Backer, J.; White, M.; Kahn, C.R. J. Biol. Chem. 266, 8302-8311, 1991  
 A;Title: Purification and partial sequence analysis of pp185, the major cellular substr  
 A;Reference number: A39811; MUID:91217066; PMID:2022647  
 A;Accession: A39811  
 A;Molecule type: protein  
 A;Residues: 44-51;173-178;223-243;489-506;635-646;871-882,'I',884,'X',886-888;932-936,'  
 A;Note: the phosphorylation residue was not identified  
 C;Comment: this protein and the beta chain of the insulin receptor itself are the major  
 C;Comment: Phosphorylation of this protein in response to insulin is maximal at 30 seco  
 C;Superfamily: insulin receptor substrate IRS-1; pleckstrin repeat homology  
 C;Keywords: phosphoprotein; signal transduction  
 F;1-113/Domain: pleckstrin repeat homology <PLK>  
 F;872-891/Region: glutamine-rich

Query Match 35.4%; Score 49.5; DB 1; Length 1235;  
 Best Local Similarity 47.8%; Pred. No. 32;  
 Matches 11; Conservative 1; Mismatches 4; Indels 7; Gaps 1;

Qy 1 CGETYSRVTH-----PHLPR 16  
 Db 816 CGARPESSVTHPHHALQPLPR 838

RESULT 13  
 B31194  
 hypothetical protein ORF1 (FGF 5' region) - human  
 C;Species: Homo sapiens (man)  
 C;Date: 12-Jan-1989 #sequence\_revision 12-Jan-1989 #text\_change 05-Nov-1999  
 C;Accession: B31194  
 R;Zhan, X.; Bates, B.; Hu, X.; Goldfarb, M.  
 Mol. Cell. Biol. 8, 3487-3495, 1988  
 A;Title: The human FGF-5 oncogene encodes a novel protein related to fibroblast growth





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 08:17:40 ; Search time 6.4 Seconds  
(without alignments)  
203.399 Million cell updates/sec

Title: US-09-701-623C-5

Perfect score: 140  
Sequence: 1 CGTYQSRVTHPLPALMRSTKC 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	117	83.6	428	1	EPC_HUMAN
2	80	57.1	429	1	EPC_RAT
3	56	40.0	421	1	EPC_MOUSE
4	50	35.7	1597	1	RRPO_CRMV
5	49.5	35.4	1235	1	IRS1_RAT
6	48	34.3	544	1	MKR3_MOUSE
7	47.5	33.9	381	1	GVPC_HALME
8	47.5	33.9	382	1	GVCI_HALN1
9	47	33.6	242	1	VMAT_INCUJ
10	47	33.6	1938	1	MYHD_HUMAN
11	47	33.6	1940	1	MYH3_CHICK
12	47	33.6	1940	1	MYH3_HUMAN
13	47	33.6	1940	1	MYH3_RAT
14	46.5	33.2	1233	1	IRS1_MOUSE
15	46	32.9	116	1	YK14_YEAST
16	46	32.9	247	1	TRY3_RAT
17	46	32.9	550	1	VGLE_HSV11
18	45	32.1	106	1	Y116_ADS02
19	45	32.1	264	1	H82D_RAT
20	45	32.1	736	1	MYH7_RABIT
21	45	32.1	777	1	BARI_HUMAN
22	45	32.1	1601	1	REPO_TVCV
23	45	32.1	1934	1	MYH7_MESAU
24	45	32.1	1935	1	MYH7_HUMAN
25	45	32.1	1935	1	MYH7_PIG
26	45	32.1	1935	1	MYH7_RAT
27	45	32.1	1937	1	MYH8_HUMAN
28	45	32.1	1938	1	MYSS_CHICK
29	45	32.1	1939	1	MYH1_HUMAN
30	45	32.1	1939	1	MYH4_HUMAN
31	45	32.1	1941	1	MYH2_HUMAN
32	45	32.1	2492	1	TALA_DICDI
33	44.5	31.8	450	1	MUC_CANFA

34 44 31.4 273 1 AROE\_YERPE  
35 44 31.4 340 1 ALC2\_HUMAN  
36 44 31.4 345 1 HA1F\_CHICK  
37 44 31.4 425 1 HE47\_CABEL  
38 44 31.4 427 1 HE47\_PIG  
39 44 31.4 428 1 HE47\_HUMAN  
40 44 31.4 428 1 HE47\_PANTR  
41 44 31.4 428 1 HE47\_RAT  
42 44 31.4 465 1 OPCA\_NOSPU  
43 44 31.4 485 1 MURE\_FUSNN  
44 44 31.4 593 1 COX1\_HALN1  
45 44 31.4 1075 1 NFC3\_HUMAN

Q82J74 yersinia pe  
P01877 homo sapien  
P15979 gallus gall  
Q18212 caenorhabdi  
Q29024 sus scrofa  
Q13838 homo sapien  
P60024 pan troglod  
Q83413 rattus norv  
P48971 nostoc punc  
Q8r635 fusobacteri  
P33518 halobacteri  
Q12968 homo sapien

#### ALIGNMENTS

RESULT 1  
EPC\_HUMAN STANDARD; PRT; 428 AA.  
AC P01554;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE IG epsilon chain C region.  
GN IGHE.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=83168897; PubMed=6300763;  
RA Seno M., Kurokawa T., Ono Y., Onda H., Sasada R., Igarashi K.,  
RA Kikuchi M., Sugino Y., Nishida Y., Honjo T.;  
RT "Molecular cloning and nucleotide sequencing of human immunoglobulin  
epsilon chain cDNA.";  
RL Nucleic Acids Res. 11:719-726(1983).  
RN [2]  
RP SEQUENCE FROM N.A., AND VARIANT LEU-359.  
RX MEDLINE=83001945; PubMed=6288268;  
RA Max E.E., Battey J., Ney R., Kirsch I.R., Leder P.;  
RT "Duplication and deletion in the human immunoglobulin epsilon genes.";  
RL Cell 29:691-699(1982).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=842336029; PubMed=6234164;  
RA Flanagan J.G., Rabbitts T.H.;  
RT "The sequence of a human immunoglobulin epsilon heavy chain constant  
region gene, and evidence for three non-allelic genes.";  
RL EMBO J. 1:655-660(1982).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=84207910; PubMed=6327276;  
RA Ueda S., Nakai S., Nishida Y., Hisajima H., Honjo T.;  
RT "Long terminal repeat-like elements flank a human immunoglobulin  
epsilon pseudogene that lacks introns.";  
RL EMBO J. 1:1539-1544(1982).  
RN [5]  
RP PRELIMINARY SEQUENCE (MYELOMA PROTEIN ND).  
RA Bennich H.H., Johansson S.G.O., von Bahr-Lindstrom H.;  
RT (In) Bach M.K. (eds.);  
RL Immediate hypersensitivity: modern concepts and developments, pp.1-36,  
Marcel Dekker, New York (1978).  
RN [6]  
RP SEQUENCE OF 1-40; 68-114 AND 427-428 FROM N.A.  
RX MEDLINE=83065234; PubMed=6815656;  
RA Kenten J.H., Molgaard H.V., Houghton M., Derbyshire R.B., Viney J.,  
RA Bell L.O., Gould H.J.;  
RT "Cloning and sequence determination of the gene for the human  
immunoglobulin epsilon chain expressed in a myeloma cell line.";  
RL Proc. Natl. Acad. Sci. U.S.A. 79:6661-6665(1982).  
RN [7]

```

RP 3D-STRUCTURE MODELING.
RX MEDLINE=87089848; PubMed=3796618;
RA Padlan E.A., Davies D.R.;
RT "A model of the Fc of immunoglobulin E.";
RL Mol. Immunol. 23:1063-1075(1986).
CC -1- SIMILARITY: Contains 4 immunoglobulin-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L00022; AAB59424.1; ALT_INIT.
DR PIR; A22771; EHHU.
DR PDB; 1IGE; 15-JUL-92.
DR PDB; 1FP5; 30-JAN-02.
DR PDB; 1G84; 16-MAY-01.
DR PDB; 1OOV; 18-SEP-02.
DR Genew; HGNC:5522; IGHE.
DR MIM; 147180; -.
DR GO; GO:0003823; F:antigen binding; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00407; IG1; 4.
DR PROSITE; PS50835; IG LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 3.
DR Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW 3D-structure; Repeat.
FT NON TER 1
FT DOMAIN 6 103 IG-LIKE 1.
FT DOMAIN 112 210 IG-LIKE 2.
FT DOMAIN 214 318 IG-LIKE 3.
FT DOMAIN 324 423 IG-LIKE 4.
FT DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 15 105
FT DISULFID 29 85
FT DISULFID 121 121 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 135 193
FT DISULFID 209 209 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 239 299
FT DISULFID 345 405
FT CARBOHYD 21 21 N-LINKED (GLCNAC. . .).
FT CARBOHYD 49 49 N-LINKED (GLCNAC. . .).
FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .).
FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .).
FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .).
FT CARBOHYD 275 275 N-LINKED (GLCNAC. . .).
FT VARIANT 359 359 W -> L (POSSIBLE POLYMORPHISM).
FT STRAND 110 110 /FTId-VAR_003985.
FT STRAND 113 117
FT HELIX 122 124
FT TURN 125 125
FT STRAND 130 141
FT STRAND 146 150
FT STRAND 155 155
FT STRAND 159 161
FT STRAND 165 166
FT TURN 168 169
FT STRAND 172 181
FT HELIX 182 186
FT TURN 187 188
FT STRAND 192 196
FT TURN 198 199
FT TURN 201 202
FT STRAND 205 207
FT STRAND 218 222

```

```

FT TURN 226 229
FT TURN 231 232
FT STRAND 236 242
FT STRAND 252 252
FT STRAND 255 257
FT TURN 258 259
FT STRAND 260 260
FT TURN 265 266
FT STRAND 271 274
FT TURN 275 276
FT STRAND 277 285
FT HELIX 288 293
FT TURN 294 294
FT STRAND 297 302
FT TURN 304 305
FT STRAND 310 314
FT STRAND 322 322
FT STRAND 325 329
FT HELIX 334 337
FT STRAND 340 348
FT STRAND 350 351
FT STRAND 356 360
FT TURN 366 369
FT STRAND 374 374
FT STRAND 377 378
FT TURN 380 381
FT STRAND 384 385
FT STRAND 387 393
FT HELIX 394 398
FT TURN 399 400
FT STRAND 404 408
FT STRAND 418 421
SQ SEQUENCE 428 AA; 47019 MW; 25C4CA072AA558A0 CRC64;

Query Match 83.6%; Score 117; DB 1; Length 428;
Best Local Similarity 95.7%; Pred. No. 2.2e-10;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GETYQSRVTHPLPALMRSTTK 24
   ||||| ||||| ||||| |||||
DB 294 GETYQSRVTHPLPALMRSTTK 316

RESULT 2
EPC_RAT
ID EPC_RAT STANDARD; PRT; 429 AA.
AC P01855;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig epsilon chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (IMMUNOCYTOMA IR2).
RC STRAIN=LOU/C/WSL;
RX MEDLINE=83064537; PubMed=6292865;
RA Hellman L., Pettersson U., Engstroem A., Karlsson T., Bennich H.;
RT "Structure and evolution of the heavy chain from rat immunoglobulin
RL Nucleic Acids Res. 10:6041-6049(1982).
RN [2]
RP SEQUENCE OF 168-342 FROM N.A. (MYELOMA IR162).
RX MEDLINE=83182019; PubMed=6820340;
RA Kindsvogel W.R., Reddy E.P., Moore J.M., Faust C.H. Jr.;
RT "A cloned cDNA probe for rat immunoglobulin epsilon heavy chain:
RL DNA 1:335-343(1982).
RN [3]
RP SEQUENCE OF 205-306 FROM N.A.
RX MEDLINE=82174576; PubMed=6803238;

```

[illegible]

```

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
-----
EMBL; X01857; CAA25977.1; -.
EMBL; X01857; CAA25978.1; -.
PIR; A02144; EHMS.
PIR; A02145; EHMSG.
DR HSSP; P01854; LIGE.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_cI.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00407; IGc1; 2.
DR PROSITE; PS0835; IG_LIKE; 4.
DR PROSITE; PS0290; IG_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1
FT DOMAIN 1 90 CH1.
FT FT DOMAIN 91 197 CH2.
FT FT DOMAIN 198 304 CH3.
FT FT DOMAIN 305 421 CH4.
FT DISULFID 23 75 BY SIMILARITY.
FT DISULFID 121 180 BY SIMILARITY.
FT DISULFID 226 285 BY SIMILARITY.
FT DISULFID 330 392 BY SIMILARITY.
FT CARBOHYD 43 43 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 95 95 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 261 261 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 365 365 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 415 415 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 421 AA; 47320 MW; 8F909E1F30A0EB47 CRC64;
Query Match 40.0%; Score 56; DB 1; Length 421;
Best Local Similarity 47.8%; Pred.No. 0.47;
Matches 11; Conservative 3; Mismatches 9; Indels 0; Gaps 0;
QY 2 GETYQSRTVTHPLFRALMRSTTK 24
DB 280 GYGQCIVDHDFPKPIVRISITK 302
RESULT 4
RRPO_CRMV STANDARD; PRG; 1597 AA.
ID ID RRPO_CRMV
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE RNA-directed RNA polymerase (EC 2.7.7.48) (182 kDa protein) [Contains:
DE Methylnitransferase/RNA helicase (MT/HEL) (125 kDa protein)].
OS Chinese rape mosaic virus (CRMV) [Oilseed rape mosaic virus].
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
OX NCBI_TaxId=42007;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=96197410; PubMed=8616237;
RX Aguilar I., Sanchez F., Martin-Martin A., Martinez-Herrera D.,
RA Ponz F.;
RA "Nucleotide sequence of Chinese rape mosaic virus (oilseed rape mosaic
RT virus), a crucifer tobamovirus infectious on Arabidopsis thaliana.";
RL Plant Mol. Biol. 30:191-197(1996).
CC -!- FUNCTION: THE LONGER PROTEIN IS A RNA POLYMERASE ACTIVE IN VIRAL
RNA REPLICATION CC

```



```
RX MEDLINE=99214019; PubMed=10196368;
RA Jong M.T.C., Carey A.H., Caldwell K.A., Lau M.H., Handel M.A.,
RA Driscoll D.J., Stewart C.L., Rinchik E.M., Nicholls R.D.;
RA "Imprinting of a RING zinc-finger encoding gene in the mouse
RT chromosome region homologous to the Prader-Willi syndrome genetic
RT region."
RL Hum. Mol. Genet. 8:795-803(1999).
CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
CC -1- SIMILARITY: Contains 3 C3H1-type zinc fingers.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U19106; AAA76863.1; -.
DR MGD; MGI:2181178; Mkrn3
DR InterPro; IPR000571; Znf_CCH.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF00097; Zf-C3HC4; 1.
DR Pfam; PF00642; Zf-CCCH; 3.
DR SMART; SM00184; RING; 1.
DR SMART; SM00356; Znf C3H1; 3.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
KW Zinc-finger; Repeat.
FT ZN FING 98 C3H1-TYPE 1.
FT ZN FING 116 C3H1-TYPE 2.
FT ZN FING 280 298 C3H1-TYPE 2.
FT DOMAIN 302 329 MAKORIN-TYPE CYS-HIS.
FT ZN FING 347 401 RING-TYPE.
FT ZN FING 436 456 C3H1-TYPE 3.
FT ZN FING 456 456 C3H1-TYPE 3.
SQ SEQUENCE 544 AA; 59444 MW; FF05B7D034C5EA9P CRC64;

Query Match 34.3%; Score 48; DB 1; Length 544;
Best Local Similarity 45.0%; Pred. No. 10;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 2 GETYQSRVTHPLRALMS 21
DB 226 GQYRGVVPFGPEAPLQS 245

RESULT 7
GVPC HALME STANDARD; PRT; 381 AA.
AC Q02228;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Gas vesicle protein C.
GN GVPC.
OS Halobacterium mediterranei (Haloflex mediterranei).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Haloflex.
OX NCBI_TaxID=2252;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 1411;
RX MEDLINE=93021102; PubMed=1404376;
RA Englart C., Krueger K., Offner S., Pfeifer F.;
RT "Three different but related gene clusters encoding gas vesicles in
RT halophilic archaea."
RL J. Mol. Biol. 227:586-592(1992).
CC -1- FUNCTION: May confer stability to the gas vesicle membranes. Gas
CC vesicles are small, hollow, gas filled protein structures that are
CC found in several microbial planktonic microorganisms. They allow
CC the positioning of the organism at the favorable depth for growth.
CC -1- SUBCELLULAR LOCATION: Binds to the external surface of the gas
CC vesicle membrane.
CC -1- SIMILARITY: VERY DISTANT, TO GAS VESICLE PROTEINS TYPE C OF
```

```
CC CYANOBACTERIA.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X64701; CAA45944.1; -.
DR PIR; S28115; S28115.
DR InterPro; IPR008639; Halo_GVPC.
DR Pfam; PF05465; Halo_GVPC; 1.
KW Gas vesicle; Repeat.
FT DOMAIN 22 274 7 X APPROXIMATE TANDEM REPEATS.
FT REPEAT 22 59 1.
FT REPEAT 22 60 2.
FT REPEAT 85 122 3.
FT REPEAT 123 160 4.
FT REPEAT 161 192 5.
FT REPEAT 193 232 6.
FT REPEAT 233 274 7.
FT DOMAIN 246 335 ASP/GLU-RICH (ACIDIC).
SQ SEQUENCE 381 AA; 42653 MW; 9FB48199D0305921 CRC64;

Query Match 33.9%; Score 47.5; DB 1; Length 381;
Best Local Similarity 64.3%; Pred. No. 8.4;
Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

OY 1 CGETYQSRVTHPL 14
DB 343 CGEYQQA-ITEPHL 355

RESULT 8
GVCI HALNI STANDARD; PRT; 382 AA.
ID GVCI HALNI
AC P24574; Q9H117;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Gas vesicle protein C.1.
GN (GVPC1 OR GVPC OR VNG5032G) AND (GVPC2 OR VNG6031G).
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081), and
OG Plasmid pNRC100, Plasmid pNRC200, and Plasmid pH1.
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091, 148370;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRC-1; PLASMID=pNRC100;
RX MEDLINE=91323716; PubMed=1864501;
RA Jones J.G., Young D.C., Dassarma S.;
RT "Structure and organization of the gas vesicle gene cluster on the
RT Halobacterium halobium plasmid pNRC100."
RL Gene 102:117-122(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NRC-1; PLASMID=pNRC100;
RX MEDLINE=95063795; PubMed=9847077;
RA Ng W.V., Ciufo S.A., Smith T.M., Bumgarner R.E., Baskin D., Faust J.,
RA Hall B., Loretz C., Seto J., Slagel J., Hood L., DasSarma S.;
RT "Snapshot of a large dynamic replicon in a halophilic archaeon:
RT megaplasmid or minichromosome?";
RL Genome Res. 8:1131-1141(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=NRC-1; PLASMID=pNRC200;
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
```

```
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welter R., Goo Y.A.,
RA Leithaus B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Idenbarger T.A., Peck R.F., Pohlenschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.,
RT "Genome sequence of Halobacterium species NRC-1.",
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=NRC-817; PLASMID=pHH1;
RX MEDLINE=92065812; PubMed=1956294;
RA Horne M., Englert C., Wimmer C., Pfeifer F.;
RT "A DNA region of 9 kbp contains all genes necessary for gas vesicle
RT synthesis in halophilic archaeobacteria.";
RL Mol. Microbiol. 5:1159-1174(1991).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=NRC-817; PLASMID=pHH1;
RX MEDLINE=93021102; PubMed=1404376;
RA Englert C., Krueger K., Offner S., Pfeifer F.;
RT "Three different but related gene clusters encoding gas vesicles in
RT halophilic archaea.";
RL J. Mol. Biol. 227:586-592(1992).
CC -!- FUNCTION: May confer stability to the gas vesicle membranes. Gas
CC vesicles are small, hollow, gas filled protein structures that are
CC found in several microbial planktonic microorganisms. They allow
CC the positioning of the organism at the favorable depth for growth.
CC -!- SUBCELLULAR LOCATION: Binds to the external surface of the gas
CC vesicle membrane.
CC -!- SIMILARITY: VERY DISTANT, TO GAS VESICLE PROTEINS TYPE C OF
CC CYANOBACTERIA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M58557; AAA98197.1; -
DR EMBL; AF016485; AAC82810.1; -
DR EMBL; AE005142; AAG20727.1; -
DR EMBL; X57161; CAA40451.1; -
DR EMBL; X64729; CAA45981.1; -
DR PIR; T08243; T08243.
DR InterPro; IPR008639; Halo.GVPC.
DR Pfam; PF05465; Halo.GVPC; 1.
KW Gas vesicle; Plasmid; Repeat; Complete proteome.
FT DOMAIN 22 284 7 X APPROXIMATE TANDEM REPEATS.
FT REPEAT 22 60 1.
FT REPEAT 61 92 2.
FT REPEAT 93 130 3.
FT REPEAT 131 168 4.
FT REPEAT 169 200 5.
FT REPEAT 201 240 6.
FT REPEAT 241 284 7.
FT DOMAIN 254 336
SQ SEQUENCE 382 AA; 42391 MW; 171DBE4C0364F46 CRC64;

Query Match 33.9%; Score 47.5; DB 1; Length 382;
Best Local Similarity 64.3%; Pred. No. 8.4;
Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

Qy 1 CGETYQSRVTHPHL 14
Db 344 CGEYQQA-ITEPHL 356

RESULT 9
VMAT_INCUJ
ID_VMAT_INCUJ STANDARD; PRT; 242 AA.

Query Match 33.9%; Score 47.5; DB 1; Length 382;
Best Local Similarity 64.3%; Pred. No. 8.4;
Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

Qy 1 CGETYQSRVTHPHL 14
Db 344 CGEYQQA-ITEPHL 356
```

```
AC P12446;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Matrix protein.
GN M.
OS Influenza C virus (strain C/Ju/50).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenzavirus C.
OX NCBI_TaxID=11560;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88300888; PubMed=3404579;
RA Yamashita M., Krystal M., Palese P.;
RT "Evidence that the matrix protein of influenza C virus is coded for
RT by a spliced mRNA.";
RL J. Virol. 62:3348-3355(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M22038; AAA43781.1; -
DR PIR; A28878; MFIVCJ.
DR InterPro; IPR004271; CM1.
DR Pfam; PF03026; CM1; 1.
KW Matrix protein.
SQ SEQUENCE 242 AA; 26967 MW; ABF9D9054E1C9D91 CRC64;

Query Match 33.6%; Score 47; DB 1; Length 242;
Best Local Similarity 38.9%; Pred. No. 6.1;
Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 CGETYQSRVTHPHLPRAL 18
Db 209 CNHTFGSNIMRPHLEKAI 226

RESULT 10
MYHD_HUMAN
ID MYHD_HUMAN STANDARD; PRT; 1938 AA.
AC Q9UKX3; O95252;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Myosin heavy chain, skeletal muscle, extraocular (MyHC-ec).
GN MYH13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Extraocular muscle;
RX MEDLINE=99318869; PubMed=10388558;
RA Weiss A., Schiaffino S., Leinwand L.A.;
RT "Comparative sequence analysis of the complete human sarcomeric myosin
RT heavy chain family: implications for functional diversity.";
RL J. Mol. Biol. 290:61-75(1999).
RN [2]
RP SEQUENCE OF 1917-1938 FROM N.A.
RC TISSUE=Extraocular muscle;
RX MEDLINE=99026150; PubMed=9806854;
RA Winters L.M., Briggs M.M., Schachat F.;
RT "The human extraocular muscle myosin heavy chain gene (MYH13) maps to
RT the cluster of fast and developmental myosin genes on chromosome 17.";
RL Genomics 54:188-189(1998).
CC -!- FUNCTION: Muscle contraction.
CC -!- SUBUNIT: Muscle myosin is a hexameric protein that consists of 2
```

heavy chain subunits (MHC), 2 alkali light chain subunits (MLC) and 2 regulatory light chain subunits (MLC-2).

-1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.

-1- DOMAIN: The rodlike tail sequence is highly repetitive, showing cycles of a 28-residue repeat pattern composed of 4 heptapeptides, characteristic for alpha-helical coiled coils.

-1- MISCELLANEOUS: Each myosin heavy chain can be split into 1 light meromyosin (LMW) and 1 heavy meromyosin (HMW). It can later be split further into 2 globular subfragments (S1) and 1 rod-shaped subfragment (S2).

-1- SIMILARITY: Contains 1 myosin-like globular head domain.

-1- SIMILARITY: Contains 1 IQ domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AF111782; AAD29948.1; --  
 EMBL; AF075248; AAC83241.1; --  
 HSP; P13538; 2MYS.  
 Genew; HGNC:7571; MYH3.  
 MIM; 603487; --  
 GO; GO:0005859; C:muscle myosin; TAS.  
 GO; GO:0003779; F:actin binding; NAS.  
 GO; GO:0005524; F:ATP binding; NAS.  
 GO; GO:0005516; F:calmodulin binding; NAS.  
 GO; GO:0003776; F:muscle motor activity; TAS.  
 GO; GO:0006936; F:muscle contraction; TAS.  
 InterPro; IPR000048; IQ region.  
 InterPro; IPR001609; myosin head.  
 InterPro; IPR004009; Myosin\_N.  
 InterPro; IPR002928; Myosin\_tail.  
 Pfam; PF00612; IQ; 2.  
 Pfam; PF00083; myosin\_head; 1.  
 Pfam; PF02736; Myosin\_N; 1.  
 PRINTS; PR00193; MYOSINHEAVY.  
 ProDom; PD000355; myosin\_head; 1.  
 SMART; SM00242; MYSC; 1.  
 SMART; PS00096; IQ; 1.  
 MYOSIN; Muscle protein; Coiled coil; Thick filament; Actin-binding; Calmodulin-binding; ATP-binding; Methylation; Multigene family.

Query Match 33.6%; Score 47; DB 1; Length 1938;  
 Best Local Similarity 52.9%; Pred No; 58;  
 Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 8 RVTHPHLPALMRSTTK 24  
 D 667 RSTHPHVRCLIPNETK 683

RESULT 11  
 MYH3\_CHICK  
 ID MYH3\_CHICK STANDARD; PRT; 1940 AA.  
 AC P02565;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Myosin heavy chain, fast skeletal muscle, embryonic.

GN MYH3.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 EX MEDLINE=87194881; PubMed=3571266;  
 RA Molina M.I., Kropp K.E., Gulick J., Robbins J.;  
 RT "The sequence of an embryonic myosin heavy chain gene and isolation  
 J. Biol. Chem. 262:6478-6488(1987).  
 RL [2]  
 RP SEQUENCE OF 1502-1940 FROM N.A.  
 EX MEDLINE=83161144; PubMed=5833296;  
 RA Kavinsky C.J., Umeda P.K., Sinha A.M., Elzinga M., Tong S.W., Zak R.,  
 RA Jakovic S., Rabinowitz M.;  
 RT "Cloned mRNA sequences for two types of embryonic myosin heavy chains  
 from chick skeletal muscle. I. DNA and derived amino acid sequence of  
 light meromyosin.";  
 RL J. Biol. Chem. 258:5196-5205(1983).  
 CC -1- FUNCTION: Muscle contraction.  
 CC -1- SUBUNIT: Muscle myosin is a hexameric protein that consists of 2  
 heavy chain subunits (MHC), 2 alkali light chain subunits (MLC)  
 and 2 regulatory light chain subunits (MLC-2).  
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
 CC -1- DOMAIN: The rodlike tail sequence is highly repetitive, showing  
 cycles of a 28-residue repeat pattern composed of 4 heptapeptides,  
 characteristic for alpha-helical coiled coils.  
 CC -1- MISCELLANEOUS: Each myosin heavy chain can be split into 1 light  
 meromyosin (LMW) and 1 heavy meromyosin (HMW). It can later be  
 split further into 2 globular subfragments (S1) and 1 rod-shaped  
 subfragment (S2).  
 CC -1- SIMILARITY: Contains 1 myosin-like globular head domain.  
 CC -1- SIMILARITY: Contains 1 IQ domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; V00430; CA23712.1; --  
 EMBL; J02714; AAA48972.1; --  
 PIR; A29320; A29320.  
 HSP; P13538; 2MYS.  
 InterPro; IPR000048; IQ region.  
 InterPro; IPR001609; myosin head.  
 InterPro; IPR004009; Myosin\_N.  
 InterPro; IPR002928; Myosin\_tail.  
 Pfam; PF00612; IQ; 2.  
 Pfam; PF00083; myosin\_head; 1.  
 Pfam; PF02736; Myosin\_N; 1.  
 PRINTS; PR00193; MYOSINHEAVY.  
 ProDom; PD000355; myosin\_head; 1.  
 SMART; SM00015; IQ; 1.  
 SMART; SM00242; MYSC; 1.  
 PROSITE; PS00096; IQ; 1.  
 MYOSIN; Muscle protein; Coiled coil; Thick filament; Actin-binding; Calmodulin-binding; ATP-binding; Methylation; Multigene family.

FT DOMAIN 1 785  
 FT DOMAIN 786 815  
 FT DOMAIN 844 1940  
 FT NP\_BIND 179 186  
 FT DOMAIN 660 682  
 FT DOMAIN 762 776  
 FT MOD\_RES 130 130  
 FT VARIANT 379 379  
 FT CONFLICT 1547 1547



```
FT CONFLICT 1913 1915 ERA -> GRT (IN REF. 2).
SQ SEQUENCE 1940 AA; 222816 MW; C348333D75B04DF2 CRC64;

Query Match 33.6%; Score 47; DB 1; Length 1940;
Best Local Similarity 52.9%; Pred. No. 58;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 8 RVTHPHLPALMRSTTK 24
Db 668 RSTHPHFVRCIIPNETK 684

RESULT 12
MYH3_HUMAN STANDARD; PRT; 1940 AA.
AC P11055; Q15492;
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Myosin heavy chain, fast skeletal muscle, embryonic (Muscle embryonic
DE myosin heavy chain) (SWHCE).
GN MYH3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89263803; PubMed=2726495;
RA Eller M.S., Stedman H.H., Sylvester J.E., Fertels S.H., Wu Q.-L.,
RA Raychowdhury M.K., Rubinstein N.A., Kelly A.M., Sarkar S.;
RA "Human embryonic myosin heavy chain cDNA. Interspecies sequence
RT conservation of the myosin rod, chromosomal locus and isoform
RT specific transcription of the gene.";
RL PDBS Lett. 256:21-28(1989).
RN [3]
RN SEQUENCE OF 856-1940 FROM N.A.
RP TISSUE=Skeletal muscle;
RX MEDLINE=90235862; PubMed=1691980;
RA Bober E., Buchberger-Seidl A., Braun T., Singh S., Goedde H.W.,
RA Arnold H.H.;
RT Identification of three developmentally controlled isoforms of human
RT myosin heavy chains.";
RL Eur. J. Biochem. 189:55-65(1990).
RN [4]
RN SEQUENCE OF 856-1940 FROM N.A.
RX MEDLINE=89366648; PubMed=2771643;
RA Karach-Mizrachi I., Travis M., Blau H., Leinwand L.A.;
RA "Expression and DNA sequence analysis of a human embryonic skeletal
RT muscle myosin heavy chain gene.";
RL Nucleic Acids Res. 17:6167-6179(1989).
CC -1- FUNCTION: Muscle contraction.
CC heavy chain subunits (MHC), 2 alkali light chain subunits (MLC)
CC and 2 regulatory light chain subunits (MLC-2).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -1- DEVELOPMENTAL STAGE: Abundantly present in fetal skeletal muscle
CC and not present or barely detectable in heart and adult skeletal
CC muscle.
CC -1- DOMAIN: The rodlike tail sequence is highly repetitive, showing
CC cycles of a 28-residue repeat pattern composed of 4 heptapeptides,
CC characteristic for alpha-helical coiled coils.
CC -1- MISCELLANEOUS: Each myosin heavy chain can be split into 1 light
CC meromyosin (LMW) and 1 heavy meromyosin (HMW). It can later be
CC split further into 2 globular subfragments (S1) and 1 rod-shaped
```

RX MEDLINE=87060988; PubMed=3783701;  
RA Strehler E.B.; Strehler-Page M.-A.; Perriard J.C., Periasamy M.,  
RT Nadal-Ginard B.;  
RA "Complete nucleotide and encoded amino acid sequence of a mammalian  
RT myosin heavy chain gene. Evidence against intron-dependent evolution  
of the rod";  
RL J. Mol. Biol. 190:291-317(1986).  
CC  
CC -1- FUNCTION: Muscle contraction.  
CC -1- SUBUNIT: Muscle myosin is a hexameric protein that consists of 2  
CC heavy chain subunits (MHC), 2 alkali light chain subunits (MLC)  
CC and 2 regulatory light chain subunits (MLC-2).  
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
CC -1- DOMAIN: The rodlike tail sequence is highly repetitive, showing  
CC cycles of a 28-residue repeat pattern composed of 4 heptapeptides,  
CC characteristic for alpha-helical coiled coils.  
CC -1- MISCELLANEOUS: Each myosin heavy chain can be split into 1 light  
CC meromyosin (LMM) and 1 heavy meromyosin (HMM). It can later be  
CC split further into 2 globular subfragments (S1) and 1 rod-shaped  
CC subfragment (S2).  
CC -1- SIMILARITY: Contains 1 myosin-like globular head domain.  
CC -1- SIMILARITY: Contains 1 IQ domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC ENBL; X04267; CAA27817.1; -;  
DR PIR; A24922; A24922.  
DR HSP; P13538; 2MYS.  
DR InterPro; IPR000048; IQ region.  
DR InterPro; IPR001603; myosin head.  
DR InterPro; IPR004009; Myosin\_N.  
DR InterPro; IPR002928; Myosin\_tail.  
DR InterPro; IPR02017; Spectrin.  
DR Pfam; PF00612; IQ; 2.  
DR Pfam; PF00063; myosin head; 1.  
DR Pfam; PF02736; Myosin\_N; 1.  
DR Pfam; PF01576; Myosin\_tail; 1.  
DR PRINTS; PR00193; MYOSINHEAVY.  
DR PRODOM; PD000355; myosin\_head; 1.  
DR SMART; SM00015; IQ; 1.  
DR SMART; SM00242; MYSC; 1.  
DR PROSITE; PS50096; IQ; 1.  
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;  
KW Calmodulin-binding; ATP-binding; Methylation; Multigene family.  
FT DOMAIN 1 781 MYOSIN HEAD-LIKE.  
FT DOMAIN 782 811 IQ.  
FT DOMAIN 840 1933 COILED COIL (POTENTIAL).  
FT NP\_BIND 179 186 ATP (POTENTIAL).  
FT DOMAIN 656 678 ACTIN-BINDING.  
FT DOMAIN 758 772 ACTIN-BINDING.  
FT MOD\_RES 130 130 METHYLATION (TRI-) (POTENTIAL).  
SQ SEQUENCE 1940 AA; 223857 MW; B5D546A596E5A696 CRC64;  
Query Match 33.6%; Score 47; DB 1; Length 1940;  
Best Local Similarity 47.1%; Pred. No. 58;  
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;  
QY 8 RVTHPHLPALMRSTTK 24  
Db 664 RTHPHFVRCILIPNETK 680  
RESULT 14  
IRS1\_MOUSE STANDARD; PRT; 1233 AA.  
ID AC P35569;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Insulin receptor substrate-1.  
GN IRS1 OR IRS-1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94220494; PubMed=8167159;  
RA Araki E., Haag B.L. III, Kahn C.R.;  
RT "Cloning of the mouse insulin receptor substrate-1 (IRS-1) gene and  
RT complete sequence of mouse IRS-1";  
RL Biochim. Biophys. Acta 1221:353-356(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93192326; PubMed=8448209;  
RA Keller S.R., Aebersold R., Garner C.W., Lienhard G.E.;  
RT "The insulin-elicited 160 kDa phosphotyrosine protein in mouse  
RT adipocytes is an insulin receptor substrate 1: identification by  
RT cloning";  
RL Biochim. Biophys. Acta 1172:323-326(1993).  
CC -1- FUNCTION: May mediate the control of various cellular processes by  
CC insulin. When phosphorylated by the insulin receptor binds  
CC specifically to various cellular proteins containing SH2 domains  
CC such as phosphatidylinositol 3-kinase p85 subunit or GRB2.  
CC -1- SIMILARITY: Contains 1 PH domain.  
CC -1- SIMILARITY: Contains 1 PTB domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC ENBL; L24563; AAA39335.1; -;  
DR EMBL; X69722; CAA49378.1; -;  
DR PIR; S30185; S30185.  
DR PDB; IAYB; 3I-AUG-94.  
DR MGD; MGI:99454; Irs1.  
DR InterPro; IPR002404; Insln\_receptorSI.  
DR InterPro; IPR001849; PH.  
DR Pfam; PF02174; IRS; 1.  
DR Pfam; PF00169; PH; 1.  
DR PRINTS; PR00628; INSULINRSI.  
DR SMART; SM00233; PH; 1.  
DR SMART; SM00310; PTB; 1.  
DR PROSITE; PS50003; PH\_DOMAIN; 1.  
KW Phosphorylation; 3D-structure.  
FT DOMAIN 12 115 PH.  
FT DOMAIN 152 262 PTB.  
FT DOMAIN 675 680 POLY-SER.  
FT DOMAIN 872 877 POLY-GLN.  
FT DOMAIN 1119 1128 POLY-GLY.  
FT DOMAIN 1194 1198 POLY-PRO.  
FT MOD\_RES 99 99 PHOSPHORYLATION (BY CK2)  
FT MOD\_RES 460 460 (BY SIMILARITY).  
FT MOD\_RES 608 608 PHOSPHORYLATION (BY INSR)  
FT MOD\_RES 608 608 (BY SIMILARITY).  
FT MOD\_RES 628 628 PHOSPHORYLATION (BY INSR)  
FT MOD\_RES 891 891 PHOSPHORYLATION (BY INSR)  
FT MOD\_RES 891 891 (BY SIMILARITY).  
FT MOD\_RES 935 935 PHOSPHORYLATION (BY INSR)  
FT MOD\_RES 983 983 PHOSPHORYLATION (BY INSR)  
FT MOD\_RES 983 983 (BY SIMILARITY).  
FT MOD\_RES 1173 1173 PHOSPHORYLATION (BY INSR)  
FT MOD\_RES 1173 1173 (BY SIMILARITY).

FT MOD\_RES 1220 1220 PHOSPHORYLATION (BY INSR)  
 (BY SIMILARITY).  
 FT CONFLICT 1038 1039 MISSING (IN REF. 2).  
 FT CONFLICT 1182 1182 H -> R (IN REF. 2).  
 SQ SEQUENCE 1233 AA; 130723 MW; COE9B2D890DADD87 CRC64;

Query Match 33.2%; Score 46.5; DB 1; Length 1233;  
 Best Local Similarity 43.5%; Pred.No. 42;  
 Matches 10; Conservative 2; Mismatches 4; Indels 7; Gaps 1;

OY 1 CGRTYQSRVTH-----PHLPR 16  
 ||| :|||  
 Db 816 CGARPESSLTHPHHVLQPHLPR 838  
 ||| :|||

RESULT 15  
 YK14 YEAST  
 ID YK14 YEAST STANDARD; PRT; 116 AA.  
 AC P36078;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE Hypothetical 13.6 kDa protein in MDH1-VMA5 intergenic region.  
 GN YKL084W.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Pohl T.M., Pohl F.M.;  
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; Z28084; CAA81922.1; -.  
 DR PIR; S37909;  
 DR Germonline; 139840; -.  
 DR SGD; S0001567; YKL084W.  
 DR InterPro; IPR008913; ZF-CHY.  
 DR Pfam; PF05495; zf-CHY; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 116 AA; 13641 MW; 57464E5FD9A591F8 CRC64;

Query Match 32.9%; Score 46; DB 1; Length 116;  
 Best Local Similarity 45.0%; Pred.No. 3.9;  
 Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

OY 6 QSRVTHPLPRALMRSTTKC 25  
 ||| :|||  
 Db 14 QSRCVHHLPKDVIAIRFKC 33  
 ||| :|||

Search completed: February 26, 2004, 08:24:18  
 Job time : 8.4 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 08:18:11 ; Search time 31.4 Seconds  
(without alignments)  
251.209 Million cell updates/sec

Title: US-09-701-623C-5

Perfect score: 140

Sequence: 1 CGETYQSRVTHPLPALMRSTTKC 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_25:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phase:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_virus:\*

16: sp\_bacteriaph:\*

17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59	42.1	532	12 Q8JNC8	Q8jnc8 macropodid
2	54	38.6	684	13 Q90544	Q90544 ginglymosto
3	53.5	38.2	333	10 Q9SY55	Q9sy55 sambucus ni
4	53.5	38.2	333	10 Q9ZT61	Q9zt61 sambucus ni
5	53.5	38.2	333	10 Q9SY54	Q9sy54 sambucus ni
6	53	37.9	505	16 Q7UL22	Q7ul22 rhodopiell
7	52	37.1	320	10 Q9LQ87	Q9lq87 arabidopsis
8	52	37.1	427	10 Q8GIW7	Q8gyw7 arabidopsis
9	52	37.1	533	16 Q9RUJ0	Q9rjy0 streptomyce
10	50.5	36.1	333	10 Q9ZT60	Q9zt60 sambucus ni
11	50	35.7	1103	12 Q91PA7	Q91pa7 ribgrass mo
12	50	35.7	1103	12 Q7TD03	Q7td03 crucifer to
13	50	35.7	1107	12 Q86503	Q86503 tobacco mcs
14	50	35.7	1197	12 Q91PA8	Q91pa8 ribgrass mo
15	50	35.7	1597	12 Q7TD02	Q7td02 crucifer to
16	50	35.7	1601	12 Q86504	Q86504 tobacco mcs

17 49.5 35.4 330 10 Q944B9  
18 49 35.0 277 10 Q9LTK6  
19 49 35.0 360 5 Q9V9X5  
20 49 35.0 840 13 Q918N3  
21 48 34.3 209 16 Q7VHN6  
22 48 34.3 1065 10 Q48948  
23 48 34.3 1065 10 Q9FHK6  
24 48 34.3 1103 12 Q8V9A1  
25 48 34.3 1103 12 Q8V9A5  
26 48 34.3 1597 12 Q8V9A2  
27 48 34.3 1597 12 Q8V9A6  
28 48 34.3 2254 10 Q9LN02  
29 47.5 33.9 2753 5 Q81KL6  
30 47 33.6 242 12 Q9WAE1  
31 47 33.6 242 12 Q91Q55  
32 47 33.6 242 12 Q67390  
33 47 33.6 242 12 Q9WAD3  
34 47 33.6 242 12 Q39842  
35 47 33.6 242 12 Q89465  
36 47 33.6 242 12 Q91Q57  
37 47 33.6 242 12 Q8AZX7  
38 47 33.6 242 12 Q8AZX5  
39 47 33.6 242 12 Q8AZX3  
40 47 33.6 242 12 Q8AZX1  
41 47 33.6 242 12 Q8AZW9  
42 47 33.6 242 12 Q8AZW7  
43 47 33.6 242 12 Q8AZW5  
44 47 33.6 242 12 Q8AZW3  
45 47 33.6 242 12 Q8AZW1

Q944b9 sambucus ni  
Q91tk6 arabidopsis  
Q9v9x5 drosophila  
Q918n3 rana pipien  
Q7vhn6 helicobacte  
Q48948 arabidopsis  
Q9fhk6 arabidopsis  
Q8v9a1 crucifer to  
Q8v9a5 crucifer to  
Q8v9a2 crucifer to  
Q8v9a6 crucifer to  
Q9ln02 arabidopsis  
Q81kl6 plasmodium  
Q9wae1 influenza c  
Q91q55 influenza c  
Q67390 influenza c  
Q9wad3 influenza c  
Q39842 influenza c  
Q89465 influenza c  
Q91q57 influenza c  
Q8azx7 influenza c  
Q8azx5 influenza c  
Q8azx3 influenza c  
Q8azx1 influenza c  
Q8azw9 influenza c  
Q8azw7 influenza c  
Q8azw5 influenza c  
Q8azw3 influenza c  
Q8azw1 influenza c

#### ALIGNMENTS

#### RESULT 1

Q8JNC8 PRELIMINARY; PRT; 532 AA.  
AC Q8JNC8;  
DT 01-OCT-2002 (Tremblrel. 22, Created)  
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Glycoprotein B.  
GN US8.  
OS Macropodid herpesvirus 1.  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Simplexvirus.  
OX NCBI\_TaxID=137443;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Thomson D., Smith G.;  
RT "Macropodid herpesvirus 1 and 2: unique short region glycoproteins."  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY048539; AAL13140.1; -  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR003404; Herpes glycopE.  
DR InterPro; IPR007110; Ig-like.  
DR Pfam; PF02480; Herpes\_5E; 1.  
SQ SEQUENCE 532 AA; 59087 MW; 567D44D52600B76B CRC64;

Query Match 42.1%; Score 59; DB 12; Length 532;

Best Local Similarity 40.7%; Pred. NO. 0.66;

Matches 11; Conservative 4; Mismatches 10; Indels 2; Gaps 1;

QY 1 CGE--TYQSRVTHPLPALMRSTTKC 25

Db 249 CGEMHYESCLHPLHPLPCTMPANAC 275

#### RESULT 2

Q90544 PRELIMINARY; PRT; 684 AA.

ID Q90544

AC Q90544

DT 01-NOV-1996 (Tremblrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Novel antigen receptor precursor.  
 OS Ginglymostoma cirratum (Nurse shark).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
 OC Elasmobranchii; Galeomorphii; Galeoidea; Orectolobiformes;  
 OC Ginglymostomatidae; Ginglymostoma.  
 CX NCBI\_TaxID=7801;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Spleen;  
 RX MEDLINE=95183140; PubMed=7877689;  
 RA Greenberg A.S., Avila D., Hughes M., Hughes A., McKinney E.C.,  
 RA Flajnik M.F.  
 RT "A new antigen receptor gene family that undergoes rearrangement and  
 RT extensive somatic diversification in sharks."  
 RL Nature 374:168-173(1995).  
 RL EMBL; U18701; AAB48195.1; -.  
 DR PIR; S60266; S60266.  
 DR HSSP; P01842; 2MCG.  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003597; Ig-cl.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig; 6.  
 DR SMART; SM00407; IG1; 4.  
 DR PROSITE; PS00835; IG\_LIKE; 6.  
 DR PROSITE; PS00290; IG\_MHC; 3.  
 KW Receptor; Signal.  
 FT SIGNAL 1 18 POTENTIAL.  
 FT CHAIN 19 684 NOVEL ANTIGEN RECEPTOR.  
 SQ SEQUENCE 684 AA; 75224 MW; 2FF9D2071CDA6DFD CRC64;  
 Query Match 38.6%; Score 54; DB 13; Length 684;  
 Best Local Similarity 57.1%; Pred. No. 5.5;  
 Matches 12; Conservative 1; Mismatches 8; Indels 0; Gaps 0;  
 QY 4 TYQSRVTHPLPALMRSTYK 24  
 DB 642 TYSCLVGHPSLNRLIRSTNK 662  
 RESULT 3  
 QSYS5  
 ID Q9SYS5 PRELIMINARY; PRT; 333 AA.  
 AC Q9SYS5;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hevein-like protein HLPf (Fragment).  
 OS Sambucus nigra (European elder).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC campanulids; Dipsacales; Adoxaceae; Sambucus.  
 CX NCBI\_TaxID=4202;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Fruit;  
 RX MEDLINE=99214486; PubMed=10198114;  
 RA Van Damme E.J.M., Charrels D., Roy S., Tiersens K., Barre A.,  
 RA Martins J.C., Rouge P., Van Leuven F., Does M., Peumans W.J.;  
 RT "A gene encoding a hevein-like protein from elderberry fruits is  
 RT homologous to PR-4 and class V chitinase genes."  
 RL Plant Physiol. 119:1547-1556(1999).  
 DR EMBL; AF074388; AAD1237.1; -.  
 DR HSSP; P02877; 1HEV.  
 DR GO; GO:0008061; F:chitin binding; IEA.  
 DR GO; GO:0004568; F:chitinase activity; IEA.  
 DR GO; GO:0016998; P:cell wall catabolism; IEA.  
 DR GO; GO:0006032; P:chitin catabolism; IEA.  
 DR GO; GO:0009613; P:response to pest/pathogen/parasite; IEA.  
 DR InterPro; IPR001002; Chitin binding\_1.  
 DR InterPro; IPR000726; Glyco\_Hydro\_19.  
 DR Pfam; PF00187; Chitin\_bind\_1; 1.  
 DR Pfam; PF00182; Glyco\_Hydro\_19; 1.  
 DR ProDom; P000609; Chitin binding\_1; 1.  
 DR ProDom; P0354900; Glyco\_Hydro\_19; 1.  
 DR SMART; SM00270; ChtBD1; 1.  
 DR PROSITE; PS00774; CHITINASE 19 2; 1.  
 DR PROSITE; PS00026; CHITIN\_BINDING; 1.  
 KW Chitin-binding.  
 SQ SEQUENCE 333 AA; 37077 MW; 3231A252FC253BES CRC64;  
 Query Match 38.2%; Score 53.5; DB 10; Length 333;  
 Best Local Similarity 42.9%; Pred. No. 3.1;  
 Matches 15; Conservative 2; Mismatches 3; Indels 15; Gaps 2;  
 QY 1 CGETVQ-----SRVTHPLPALMRST 22  
 DB 52 CGSTQYCEGCGQCRDTSLT--DLPRALLRPT 84  
 RESULT 5  
 QSYS4  
 ID Q9SYS4 PRELIMINARY; PRT; 333 AA.

DR Pfam; PF00187; Chitin\_bind\_1; 1.  
 DR Pfam; PF00182; Glyco\_Hydro\_19; 1.  
 DR ProDom; P000609; Chitin binding\_1; 1.  
 DR ProDom; P0354900; Glyco\_Hydro\_19; 1.  
 DR SMART; SM00270; ChtBD1; 1.  
 DR PROSITE; PS00774; CHITINASE 19 2; 1.  
 DR PROSITE; PS00026; CHITIN\_BINDING; 1.  
 KW Chitin-binding.  
 FT NON\_TER 333  
 SQ SEQUENCE 333 AA; 37047 MW; 3913256EBC3577B9 CRC64;  
 Query Match 38.2%; Score 53.5; DB 10; Length 333;  
 Best Local Similarity 42.9%; Pred. No. 3.1;  
 Matches 15; Conservative 2; Mismatches 3; Indels 15; Gaps 2;  
 QY 1 CGETVQ-----SRVTHPLPALMRST 22  
 DB 52 CGSTQYCEGCGQCRDTSLT--DLPRALLRPT 84  
 RESULT 4  
 Q9ZT61  
 ID Q9ZT61 PRELIMINARY; PRT; 333 AA.  
 AC Q9ZT61;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hevein-like protein.  
 OS Sambucus nigra (European elder).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC campanulids; Dipsacales; Adoxaceae; Sambucus.  
 CX NCBI\_TaxID=4202;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Fruit;  
 RX MEDLINE=99214486; PubMed=10198114;  
 RA Van Damme E.J.M., Charrels D., Roy S., Tiersens K., Barre A.,  
 RA Martins J.C., Rouge P., Van Leuven F., Does M., Peumans W.J.;  
 RT "A gene encoding a hevein-like protein from elderberry fruits is  
 RT homologous to PR-4 and class V chitinase genes."  
 RL Plant Physiol. 119:1547-1556(1999).  
 DR EMBL; AF074385; AAD11406.1; -.  
 DR HSSP; P02877; 1HEV.  
 DR GO; GO:0008061; F:chitin binding; IEA.  
 DR GO; GO:0004568; F:chitinase activity; IEA.  
 DR GO; GO:0016998; P:cell wall catabolism; IEA.  
 DR GO; GO:0006032; P:chitin catabolism; IEA.  
 DR GO; GO:0009613; P:response to pest/pathogen/parasite; IEA.  
 DR InterPro; IPR001002; Chitin binding\_1.  
 DR InterPro; IPR000726; Glyco\_Hydro\_19.  
 DR Pfam; PF00187; Chitin\_bind\_1; 1.  
 DR Pfam; PF00182; Glyco\_Hydro\_19; 1.  
 DR ProDom; P000609; Chitin binding\_1; 1.  
 DR ProDom; P0354900; Glyco\_Hydro\_19; 1.  
 DR SMART; SM00270; ChtBD1; 1.  
 DR PROSITE; PS00774; CHITINASE 19 2; 1.  
 DR PROSITE; PS00026; CHITIN\_BINDING; 1.  
 KW Chitin-binding.  
 SQ SEQUENCE 333 AA; 37077 MW; 3231A252FC253BES CRC64;  
 Query Match 38.2%; Score 53.5; DB 10; Length 333;  
 Best Local Similarity 42.9%; Pred. No. 3.1;  
 Matches 15; Conservative 2; Mismatches 3; Indels 15; Gaps 2;  
 QY 1 CGETVQ-----SRVTHPLPALMRST 22  
 DB 52 CGSTQYCEGCGQCRDTSLT--DLPRALLRPT 84  
 RESULT 5  
 QSYS4  
 ID Q9SYS4 PRELIMINARY; PRT; 333 AA.

Q9SY34;  
 01-MAY-2000 (TrEMBLrel. 13, Created)  
 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Havein-like protein.  
 OS Sambucus nigra (European elder).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC campanulids; Dipsacales; Adoxaceae; Sambucus.  
 OX NCBI\_TaxID=4202;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Fruit;  
 RX MEDLINE=99214486; PubMed=10198114;  
 RA Van Damme E.J.M., Chavels D., Roy S., Tiersens K., Barre A.,  
 RA Martins J.C., Rouge P., Van Leuven F., Does M., Peumans W.J.;  
 RT "A gene encoding a hevein-like protein from elderberry fruits is  
 RT homologous to PR-4 and class V chitinase genes";  
 RL Plant Physiol. 119:1547-1556 (1999).  
 DR EMBL: AF074367; AAD11408.1; -;  
 DR HSP; P02877; IHEV.  
 DR GO: GO:0008061; F:Chitin binding; IEA.  
 DR GO: GO:0004568; F:Chitinase activity; IEA.  
 DR GO: GO:0016998; P:Cell wall catabolism; IEA.  
 DR GO: GO:0006032; P:chitin catabolism; IEA.  
 DR GO: GO:0009613; P:response to pest/pathogen/parasite; IEA.  
 DR InterPro: IPR001002; Chitin binding\_1.  
 DR InterPro: IPR000726; Glyco\_Hydro\_19.  
 DR Pfam: PF00187; Chitin bind\_1; 1.  
 DR Pfam: PF00182; Glyco\_Hydro\_19; 1.  
 DR ProDom: PD000609; Chitin binding\_1; 1.  
 DR ProDom: PD354900; Glyco\_Hydro\_19; 1.  
 DR SMART: SM00270; CatBD1\_1.  
 DR PROSITE: PS00774; CHITINASE\_19; 2; 1.  
 DR PROSITE: PS00026; CHITIN\_BINDING; 1.  
 KW Chitin-binding.  
 SQ SEQUENCE 333 AA; 37126 MW; 3231BD42F4352CE5 CRC64;  
 Query Match 38.2%; Score 53.5; DB 10; Length 333;  
 Best Local Similarity 42.9%; Pred. No. 3.1;  
 Matches 15; Conservative 2; Mismatches 3; Indels 15; Gaps 2;  
 QY 1 CGTYQ-----SRVTHPLPALMRST 22  
 DB 52 CGSTQYQCEGQSCQCRDTSRLT--DLPRALLRPT 84  
 RESULT 6  
 QYULZ2 PRELIMINARY; PRT; 505 AA.  
 AC QYULZ2;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 GN RB9180.  
 OS Rhodospirillum baltica.  
 OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;  
 OC Planctomycetaceae; Pirellula.  
 OX NCBI\_TaxID=117;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1;  
 RX MEDLINE=22735913; PubMed=12835416;  
 RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,  
 RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,  
 RA Schlesner H., Anann R., Reinhardt R.;  
 RT "Complete genome sequence of the marine planctomycete Pirellula sp.  
 RT strain 1";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303 (2003).  
 DR EMBL: BX294149; CAD76125.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 505 AA; 54499 MW; 9041B06455C20C6F CRC64;

Query Match 37.9%; Score 53; DB 16; Length 505;  
 Best Local Similarity 50.0%; Pred. No. 5.8;  
 Matches 11; Conservative 2; Mismatches 9; Indels 0; Gaps 0;  
 QY 2 GETYQSRVTHPLPALMRSTT 23  
 DB 477 GRATQSRVDPSPMPTAPARSTT 498  
 RESULT 7  
 QYLO87 PRELIMINARY; PRT; 320 AA.  
 AC QYLO87;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE TLN6.10 protein.  
 GN TLN6.10.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia.  
 RA Liu S., Chan A., Yu G., Lee J., Lenz C., Pham P., Sakano H.,  
 RA Toriumi M., Chin C., Chiu J., Choi E., Chung M., Gonzalez A.,  
 RA Hong B., Liu A., Vaysberg M., Altati H., Brooks S., Buehler E.,  
 RA Chao Q., Conn L., Conway A., Hansen N., Johnson-Hopson C., Khan S.,  
 RA Kim C., Lam B., Miranda M., Nguyen M., Palm C., Shinn P.,  
 RA Southwick A., Davis R., Ecker J., Federspiel N., Theologis A.;  
 RT "The sequence of BAC TLN6 from Arabidopsis thaliana chromosome 1";  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC009273; AAF78401.1; -;  
 DR PIR: C86148; C86148.  
 DR HSP; P23911; 1C8U.  
 DR GO: GO:0016291; P:acyl-CoA thioesterase activity; IEA.  
 DR GO: GO:0006637; P:acyl-CoA metabolism; IEA.  
 DR InterPro: IPR003703; Acyl-CoA\_thio.  
 DR Pfam: PF02551; Acyl-CoA\_thio; 2.  
 SQ SEQUENCE 320 AA; 36157 MW; 03A82A96333B1BFC CRC64;  
 Query Match 37.1%; Score 52; DB 10; Length 320;  
 Best Local Similarity 60.0%; Pred. No. 5.2;  
 Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 QY 3 ETYQSRVTHPLPALMR 17  
 DB 142 ELRESITDPLPR 156  
 RESULT 8  
 QYGYW7 PRELIMINARY; PRT; 427 AA.  
 AC QYGYW7;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Putative acyl CoA thioesterase (At1g01710).  
 GN AT1G01710/TLN6.21.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,  
 RA Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,  
 RA Hayashizaki Y., Shinozaki K.;

```
RT "Arabidopsis thaliana full-length cDNA.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Bowser L., Carninci P.,
RA Chan M.M., Chang C.H., Dale J.M., Hayashizaki Y., Huan V.W.,
RA Ishida J., Jones T., Kamiya A., Karlun-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G., Yuan S.,
RA Shinozaki K., Davis R.M., Theologis A., Ecker J.R.;
RT "Arabidopsis ORF clones.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK117346; BA042016.1; -.
DR EMBL; BT005354; AA063418.1; -.
DR GO; GO:0016291; F:acyl-CoA thioesterase activity; IEA.
DR GO; GO:0006637; P:acyl-CoA metabolism; IEA.
DR InterPro; IPR003703; Acyl CoA thio.
DR InterPro; IPR000595; CNMP_binding.
DR Pfam; PF02551; Acyl CoA thio; 2.
DR Pfam; PF00027; CNMP_binding; 1.
DR SMART; SM00100; CNMP; 1.
DR TIGRFAMs; TIGR00189; tesB; 1.
DR PROSITE; PS0042; CNMP_BINDING_3; 1.
SQ SEQUENCE 427 AA; 48129 MW; 3254604D94333466 CRC64;

Query Match 37.1%; Score 52; DB 10; Length 427;
Best Local Similarity 60.0%; Pred. No. 7;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 ETQSRVTHPLPRA 17
DB 262 ELRESRITDPLRPS 276
| : : : : :
| : : : : :

RESULT 9
Q9RJY0
ID Q9RJY0 PRELIMINARY; PRT; 533 AA.
AC Q9RJY0
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative export protein.
GN SC01190 OR SCG11A.21C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1302;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Oliver K., Harris D.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
EX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapite D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,

RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC 1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL; AL939108; CAB61602.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR004638; Efflux_EmrB.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR Pfam; PF00083; sugar_tr_1.
DR TIGRFAMs; TIGR00711; efflux_EmrB; 1.
DR PROSITE; PS00850; MFS; 1.
DR Transmembrane; Complete proteome.
SQ SEQUENCE 533 AA; 54474 MW; 49C650B9797FEF77 CRC64;

Query Match 37.1%; Score 52; DB 16; Length 533;
Best Local Similarity 58.8%; Pred. No. 8.8;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 5 YQSRVTHPLPALMRS 21
DB 268 YETRVHEPLLEMLFRS 284
| : : : : :
| : : : : :

RESULT 10
Q9ZT60
ID Q9ZT60 PRELIMINARY; PRT; 333 AA.
AC Q9ZT60
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hevein-like protein.
OS Sambucus nigra (European elder).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Dipsacales; Adoxaceae; Sambucus.
OX NCBI_TaxID=4202;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fruit;
RX MEDLINE=99214486; PubMed=10198114;
RA Van Damme E.J.M., Charels D., Roy S., Tiersen K., Barre A.,
RA Martins J.C., Rouge P., Van Leuven F., Does M., Peumans W.J.;
RT "A gene encoding a hevein-like protein from elderberry fruits is
RT homologous to PR-4 and class V chitinase genes.";
RL Plant Physiol. 119:1547-1556(1999).
DR EMBL; AF04386; AAD11407.1; -.
DR HSP; P02877; IHEV.
DR GO; GO:0008061; F:chitin binding; IEA.
DR GO; GO:0004568; P:chitinase activity; IEA.
DR GO; GO:0016998; P:cell wall catabolism; IEA.
DR GO; GO:0006032; P:chitin catabolism; IEA.
DR GO; GO:0009613; P:response to pest/pathogen/parasite; IEA.
DR InterPro; IPR001002; Chitin_binding_1.
DR InterPro; IPR000726; Glyco_hydro_19.
DR Pfam; PF00187; Chitin_bind_1; 1.
DR Pfam; PF00182; Glyco_hydro_19; 1.
DR ProDom; PD000609; Chitin_binding_1; 1.
DR ProDom; PD354900; Glyco_hydro_19; 1.
DR SMART; SM00270; ChBD1_1.
DR PROSITE; PS00774; CHITINASE_19_2; 1.
DR PROSITE; PS00026; CHITIN_BINDING; 1.
KW Chitin-binding.
SQ SEQUENCE 333 AA; 36922 MW; 70C6969392A1174A CRC64;
```





```
ID Q91PAB PRELIMINARY; PRT; 1597 AA.
AC Q91PAB;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RNA replicase read-through component.
OS Ribgrass mosaic virus (RMV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
CX NCBI_TaxID=51680;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Shanghai;
RX MEDLINE=21395181; PubMed=11504428;
RA Zhu H., Hong J., Ye R., Chen J., Yu S., Adams M.J.;
RT "Sequence analysis shows that Ribgrass mosaic virus Shanghai isolate (RMV-Sh) is closely related to Youcai mosaic virus."
RL Arch. Virol. 146:1231-1238(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Shanghai;
RA Zhu H., Hong J., Chen J., Yu S.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF254924; AAK53541.1;
DR GO; GO:0005509; P:calcium ion binding; IEA.
DR GO; GO:0008174; P:mRNA methyltransferase activity; IEA.
DR GO; GO:0003724; P:RNA helicase activity; IEA.
DR GO; GO:0003968; P:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0006396; P:RNA processing; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR001788; RNA dep RNAPol2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR InterPro; IPR000606; Viral_helicase.
DR InterPro; IPR002588; V_methyltransf.
DR Pfam; PF00978; RNA_dep RNAPol2; 1.
DR Pfam; PF01443; Viral_helicase1; 1.
DR Pfam; PF01660; Vmethyltransf; 1.
DR PROSITE; PS00018; EF_HAND; 1.
SQ SEQUENCE 1597 AA; 181462 MW; 9B6576FF595B0DEF CRC64;

Query Match 35.7%; Score 50; DB 12; Length 1597;
Best Local Similarity 36.8%; Pred.No. 58;
Matches 14; Conservative 4; Mismatches 6; Indels 14; Gaps 2;

QY 2 GETYQS---RVT-----HPHLPRALMRSTTKC 25
DB 1031 GETYEKTAIVRLTATPLEIISRASPHVLVALTRHTTTC 1068

RESULT 15
Q7TD02 PRELIMINARY; PRT; 1597 AA.
AC Q7TD02;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RNA replicase, read-through component.
OS Crucifer tobamovirus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
CX NCBI_TaxID=78276;
RN [1]
RP SEQUENCE FROM N.A.
RA Ryabov E.V., Walsh J.;
RT "Infectious cDNA clone of Crucifer tobamovirus.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY318866; AAP79445.1;
SQ SEQUENCE 1597 AA; 181405 MW; AEB67B7779AF1807 CRC64;

Query Match 35.7%; Score 50; DB 12; Length 1597;
Best Local Similarity 36.8%; Pred.No. 58;
Matches 14; Conservative 4; Mismatches 6; Indels 14; Gaps 2;
```

```
QY 2 GETYQS---RVT-----HPHLPRALMRSTTKC 25
DB 1031 GETYEKTAIVRLTATPLEIISRASPHVLVALTRHTTTC 1068

Search completed: February 26, 2004, 08:27:10
Job time : 35.4 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 08:27:32 ; Search time 26.2 Seconds

(without alignments)  
201.482 Million cell updates/sec

Title: US-09-701-623c-5

Perfect score: 140

Sequence: 1 CGETYQSRVTHPLPALMRSTTKC 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications\_AA.\*  
1: /cgn2\_6/prodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/prodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/prodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/prodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/prodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/prodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/prodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/prodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/prodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/prodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/prodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/prodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/prodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/prodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/prodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/prodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/prodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/prodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	140	100.0	44	14	US-10-076-674-10
2	140	100.0	44	15	US-10-355-161A-10
3	140	100.0	45	14	US-10-076-674-11
4	140	100.0	45	15	US-10-355-161A-11
5	117	83.6	115	14	US-10-152-190-4
6	117	83.6	129	14	US-10-152-190-6
7	117	83.6	222	9	US-09-809-746-2
8	117	83.6	222	10	US-09-809-715-6
9	117	83.6	320	10	US-09-847-208-6
10	117	83.6	323	9	US-09-949-375A-2
11	117	83.6	323	9	US-09-949-375A-4
12	117	83.6	323	9	US-09-949-375A-6
13	117	83.6	323	9	US-09-949-375A-10
14	117	83.6	331	9	US-09-401-636-1
15	117	83.6	331	14	US-10-176-664-1

16	117	83.6	331	14	US-10-207-655-329	Sequence 329, Appl
17	117	83.6	336	9	US-09-949-375A-8	Sequence 8, Appl
18	117	83.6	342	9	US-09-401-636-8	Sequence 8, Appl
19	117	83.6	342	14	US-10-176-664-8	Sequence 8, Appl
20	117	83.6	347	14	US-10-152-190-13	Sequence 13, Appl
21	117	83.6	427	10	US-09-847-208-5	Sequence 5, Appl
22	117	83.6	428	9	US-09-916-230-1	Sequence 1, Appl
23	117	83.6	428	9	US-09-949-375A-1	Sequence 1, Appl
24	117	83.6	428	13	US-10-047-542-60	Sequence 60, Appl
25	117	83.6	441	9	US-09-949-375A-7	Sequence 7, Appl
26	117	83.6	569	10	US-09-847-208-7	Sequence 45, Appl
27	117	83.6	574	13	US-10-047-542-45	Sequence 37, Appl
28	117	83.6	574	14	US-10-214-524-37	Sequence 176, App
29	117	83.6	574	14	US-10-050-903-176	Sequence 176, App
30	117	83.6	574	14	US-10-050-898-176	Sequence 176, App
31	117	83.6	592	14	US-10-207-655-334	Sequence 334, App
32	113	80.7	426	14	US-10-214-524-26	Sequence 26, Appl
33	99.5	71.1	109	9	US-09-802-077-1	Sequence 1, Appl
34	99.5	71.1	109	9	US-09-802-096-1	Sequence 1, Appl
35	99.5	71.1	109	10	US-09-925-179-1	Sequence 33, Appl
36	93	66.4	25	9	US-09-974-449-33	Sequence 98, Appl
37	92	65.7	18	14	US-10-304-443-98	Sequence 110, App
38	92	65.7	20	14	US-10-304-443-110	Sequence 121, App
39	91	65.0	18	14	US-10-304-443-121	Sequence 116, App
40	91	65.0	20	14	US-10-304-443-116	Sequence 11, Appl
41	90	64.3	26	9	US-09-938-700-11	Sequence 1, Appl
42	85	60.7	114	14	US-10-152-190-1	Sequence 3, Appl
43	85	60.7	115	14	US-10-152-190-2	Sequence 2, Appl
44	85	60.7	117	14	US-10-152-190-3	Sequence 11, Appl
45	85	60.7	341	9	US-09-401-636-11	

## ALIGNMENTS

RESULT 1  
US-10-076-674-10  
; Sequence 10, Application US/10076674  
; Publication No. US20030165478A1  
; GENERAL INFORMATION:  
; APPLICANT: Sokoll, Kenneth K.  
; TITLE OF INVENTION: Stabilized Synthetic Immunogen Delivery System  
; FILE REFERENCE: Immunogen Delivery System  
; CURRENT APPLICATION NUMBER: US/10/076,674  
; CURRENT FILING DATE: 2002-04-23  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 10  
; LENGTH: 44  
; TYPE: PRT  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (19)..(19)  
; OTHER INFORMATION: Xaa indicates epsilon-Lys  
US-10-076-674-10

Query Match 100.0%; Score 140; DB 14; Length 44;  
Best Local Similarity 100.0%; Pred. No. 5.7e-13;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYQSRVTHPLPALMRSTTKC 25  
Db 20 CGETYQSRVTHPLPALMRSTTKC 44

RESULT 2  
US-10-355-161A-10  
; Sequence 10, Application US/1035516A  
; Publication No. US20040009897A1  
; GENERAL INFORMATION:  
; APPLICANT: Sokoll, Kenneth K.  
; TITLE OF INVENTION: Stabilized Synthetic Immunogen Delivery System

```
; FILE REFERENCE: Immunogen Delivery System
; CURRENT APPLICATION NUMBER: US/10/355,161A
; CURRENT FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: US 10/076674
; PRIOR FILING DATE: 2002-02-14
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (19)..(19)
; OTHER INFORMATION: Xaa indicates epsilon-Lys
US-10-355-161A-10

Query Match      100.0%; Score 140; DB 15; Length 44;
Best Local Similarity 100.0%; Pred. No. 5.7e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CGETYQSRVTHPHLPALMRSTTK 25
Db      20 CGETYQSRVTHPHLPALMRSTTK 44

RESULT 3
US-10-076-674-11
; Sequence 11, Application US/10076674
; Publication No. US20030165478A1
; GENERAL INFORMATION:
; APPLICANT: Sokoll, Kenneth K.
; TITLE OF INVENTION: Stabilized Synthetic Immunogen Delivery System
; FILE REFERENCE: Immunogen Delivery System
; CURRENT APPLICATION NUMBER: US/10/076,674
; CURRENT FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (20)..(20)
; OTHER INFORMATION: Xaa indicates epsilon-Lys
US-10-076-674-11

Query Match      100.0%; Score 140; DB 14; Length 45;
Best Local Similarity 100.0%; Pred. No. 5.8e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CGETYQSRVTHPHLPALMRSTTK 25
Db      21 CGETYQSRVTHPHLPALMRSTTK 45

RESULT 4
US-10-355-161A-11
; Sequence 11, Application US/10355161A
; Publication No. US20040009897A1
; GENERAL INFORMATION:
; APPLICANT: Sokoll, Kenneth K.
; TITLE OF INVENTION: Stabilized Synthetic Immunogen Delivery System
; FILE REFERENCE: Immunogen Delivery System
; CURRENT APPLICATION NUMBER: US/10/355,161A
; CURRENT FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: US 10/076674
; PRIOR FILING DATE: 2002-02-14
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (20)..(20)
; OTHER INFORMATION: Xaa indicates epsilon-Lys
US-10-076-674-11

Query Match      100.0%; Score 140; DB 14; Length 45;
Best Local Similarity 100.0%; Pred. No. 5.8e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 CGETYQSRVTHPHLPALMRSTTK 25
Db      21 CGETYQSRVTHPHLPALMRSTTK 45

RESULT 5
US-10-152-190-4
; Sequence 4, Application US/10152190
; Publication No. US20030096369A1
; GENERAL INFORMATION:
; APPLICANT: Morsey, Mohamad A.
; TITLE OF INVENTION: No. US20030096369A1-anaphylactogenic IgE vaccines
; FILE REFERENCE: PC11011A
; CURRENT APPLICATION NUMBER: US/10/152,190
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Human CH3
US-10-152-190-4

Query Match      83.6%; Score 117; DB 14; Length 115;
Best Local Similarity 95.7%; Pred. No. 3.1e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 GETYQSRVTHPHLPALMRSTTK 24
Db      85 GETYQSRVTHPHLPALMRSTTK 107

RESULT 6
US-10-152-190-6
; Sequence 6, Application US/10152190
; Publication No. US20030096369A1
; GENERAL INFORMATION:
; APPLICANT: Morsey, Mohamad A.
; TITLE OF INVENTION: No. US20030096369A1-anaphylactogenic IgE vaccines
; FILE REFERENCE: PC11011A
; CURRENT APPLICATION NUMBER: US/10/152,190
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Baculovirus expressed human CH3 domain
US-10-152-190-6

Query Match      83.6%; Score 117; DB 14; Length 129;
Best Local Similarity 95.7%; Pred. No. 3.5e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 GETYQSRVTHPHLPALMRSTTK 24
Db      105 GETYQSRVTHPHLPALMRSTTK 127

RESULT 7
US-09-809-746-2
```

```
; Sequence 2, Application US/09809746
; Patent No. US20010039479A1
; GENERAL INFORMATION:
; APPLICANT: Jardetzky, Theodore S.
; APPLICANT: Wurzburg, Beth A.
; TITLE OF INVENTION: THREE-DIMENSIONAL MODEL OF A FC REGION OF AN Ige
; TITLE OF INVENTION: ANTIBODY AND USES THEREOF
; FILE REFERENCE: AL-9-C2
; CURRENT APPLICATION NUMBER: US/09/809,746
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/234,877
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/189,403
; PRIOR FILING DATE: 2000-03-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-809-746-2

Query Match      83.6%; Score 117; DB 9; Length 222;
Best Local Similarity 95.7%; Pred. No. 6.2e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 GETYQSRVTHPHLPALMRSTTK 24
Db      88 GETYQCRVTHPHLPALMRSTTK 110

RESULT 8
US-09-809-715-6
; Sequence 6, Application US/09809715
; Publication No. US20030003502A1
; GENERAL INFORMATION:
; APPLICANT: Jardetzky, Theodore S.
; APPLICANT: Garman, Scott Clayton
; APPLICANT: Wurzburg, Beth A.
; APPLICANT: Kinet, Jean-Pierre
; TITLE OF INVENTION: THREE-DIMENSIONAL MODEL OF A COMPLEX BETWEEN A FC
; TITLE OF INVENTION: EPSILON RECEPTOR ALPHA CHAIN AND A FC REGION OF AN Ige
; TITLE OF INVENTION: ANTIBODY AND USES THEREOF
; FILE REFERENCE: AL-8
; CURRENT APPLICATION NUMBER: US/09/809,715
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/189,853
; PRIOR FILING DATE: 2000-03-15
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-809-715-6

Query Match      83.6%; Score 117; DB 10; Length 222;
Best Local Similarity 95.7%; Pred. No. 6.2e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 GETYQSRVTHPHLPALMRSTTK 24
Db      88 GETYQCRVTHPHLPALMRSTTK 110

RESULT 9
US-09-847-208-6
; Sequence 6, Application US/09847208
; Publication No. US20030082190A1
; GENERAL INFORMATION:
; APPLICANT: Saxon, Andrew
; APPLICANT: Zhang, Ke
; APPLICANT: Zhu, Daocheng
; TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF
; TITLE OF INVENTION: IGE-MEDIATED ALLERGIC DISEASES
; FILE REFERENCE: UC67.002A
; CURRENT APPLICATION NUMBER: US/09/847,208
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-847-208-6

Query Match      83.6%; Score 117; DB 10; Length 320;
Best Local Similarity 95.7%; Pred. No. 9e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 GETYQSRVTHPHLPALMRSTTK 24
Db      186 GETYQCRVTHPHLPALMRSTTK 208

RESULT 10
US-09-949-375A-2
; Sequence 2, Application US/09949375A
; Patent No. US20020172673A1
; GENERAL INFORMATION:
; APPLICANT: KLYSNER, Steen et al.
; TITLE OF INVENTION: METHOD FOR DOWN-REGULATING IGE
; FILE REFERENCE: 3631-0111P
; CURRENT APPLICATION NUMBER: US/09/949,375A
; CURRENT FILING DATE: 2002-01-18
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 323
; TYPE: PRT
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (8)..(103)
; OTHER INFORMATION: Human Ige heavy chain C2 domain
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (112)..(211)
; OTHER INFORMATION: Human Ige heavy chain C3 domain
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (216)..(317)
; OTHER INFORMATION: Human Ige heavy chain C4 domain
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (104)..(111)
; OTHER INFORMATION: Linker between domains C2 and C3
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (212)..(215)
; OTHER INFORMATION: Linker between domains C3 and C4
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (100)..(114)
; OTHER INFORMATION: Epitope including C2C3 linker
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (210)..(218)
; OTHER INFORMATION: Epitope including C3C4 linker
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (139)..(145)
; OTHER INFORMATION: Epitope in BC loop
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (167)..(175)
```



```

; PRIOR APPLICATION NUMBER: US 60/106,652
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated proteins
US-10-176-664-1

Query Match      83.6%; Score 117; DB 14; Length 331;
Best Local Similarity 95.7%; Pred.No. 9.3e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 GETYQSRVTHPHLPALMRSTTK 24
      ||||| ||||| ||||| |||||
Db      197 GETYQCRVTHPHLPALMRSTTK 219

```

Search completed: February 26, 2004, 08:46:53  
Job time : 27.2 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 08:18:56 ; Search time 10 Seconds  
(without alignments)  
240.479 Million cell updates/sec

Title: US-09-701-623C-6

Perfect score: 141

Sequence: 1 GETYYSRVTHPLPKDIVRSIAKC 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	88	62.4	426	2 I36948	Ig epsilon-chain - chimpanzee
2	85	60.3	428	2 EHHU	Ig epsilon chain C
3	82	58.2	429	1 EHRT	Ig epsilon chain C
4	58	41.1	107	2 I68730	IgE chain C3 regio
5	58	41.1	107	2 I68726	IgE chain C3 regio
6	58	41.1	388	1 EHMS	Ig epsilon chain C
7	58	41.1	548	2 S38864	Ig epsilon chain C
8	54	38.3	116	2 S37909	hypothetical prote
9	54	38.3	394	2 T19571	hypothetical prote
10	54	38.3	1306	2 S42659	acetyl-CoA carboxy
11	54	38.3	1625	2 T02921	acetyl-CoA carboxy
12	54	38.3	2311	2 T06161	acetyl-CoA carboxy
13	54	38.3	2325	2 T02235	acetyl-CoA carboxy
14	52	36.9	1756	2 T02599	hypothetical prote
15	51.5	36.5	605	1 FPMS	alpha-fetoprotein
16	51	36.2	106	2 A03859	hypothetical prote
17	51	36.2	195	2 A75171	hypothetical prote
18	50	35.5	539	2 T48197	hypothetical prote
19	49.5	35.1	381	2 S28115	gas-vesicle protei
20	49.5	35.1	382	2 T01222	gas-vesicle protei
21	49.5	35.1	382	2 T08243	gas-vesicle operon
22	49	34.8	684	2 S60266	novel antigen rece
23	49	34.8	1095	2 C59437	KIAA0456 protein [
24	49	34.8	1685	2 T02750	acetyl-CoA carboxy
25	48	34.0	367	2 D97096	probable permease
26	48	34.0	384	2 T06814	GTP-binding regula
27	48	34.0	454	2 A46532	Ig mu chain C regi
28	48	34.0	958	2 S15566	pol protein - simi
29	48	34.0	1161	2 S18738	pol protein - simi

30 47 33.3 110 2 S43147 Ig upsilon chain -  
31 47 33.3 320 2 C86148 hypothetical prote  
32 47 33.3 352 2 A80595 molybdenum transpo  
33 47 33.3 372 2 T16971 probable GRP-bindin  
34 47 33.3 384 1 RGT00A GTP-binding regula  
35 47 33.3 2254 2 D86215 protein T6D22.14 [I  
36 47 33.3 2257 2 T09538 acetyl-CoA carboxy  
37 46.5 33.0 242 1 C97419 probable phosphoe  
38 46.5 33.0 242 2 A28637 conserved hypochet  
39 46.5 33.0 248 2 T48376 hypothetical prote  
40 46 32.6 198 2 E71044 hypothetical prote  
41 46 32.6 245 2 I55951 MHC class II E-bet  
42 46 32.6 264 2 A60497 H-2 class II histo  
43 46 32.6 264 2 S10989 class II histocomp  
44 46 32.6 382 2 T34840 probable membrane  
45 46 32.6 549 2 S04845 Ig heavy chain pre

#### ALIGNMENTS

##### RESULT 1

I36948

Ig epsilon-chain - chimpanzee (fragment)

C:Species: Pan troglodytes (chimpanzee)

C>Date: 04-Oct-1996 #sequence\_revision 04-Oct-1996 #text\_change 21-Jan-2000

C/Accession: I36948

R:Sakoyama, Y.; Hong, K.

Proc. Natl. Acad. Sci. U.S.A. 84, 1080-1084, 1987

A:Title: Nucleotide sequences of immunoglobulin epsilon genes of chimpanzee and orangut.

A:Reference number: I36948; MUID:87147196; PMID:3103123

A/Accession: I36948

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-426 <RSS>

A/Cross-references: GB:M15398; NID:G176797; PIDN:AAA35416.1; PID:G176798

C/Genetics:

A:Introns: 103/1; 209/1; 317/1

C:Superfamily: immunoglobulin C region; immunoglobulin homology

F:336-405/Domain: immunoglobulin homology <IMM>

Query Match 62.4%; Score 88; DB 2; Length 426;

Best Local Similarity 69.6%; Pred. No. 4.8e-05;

Matches 16; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 GETYYSRVTHPLPKDIVRSIAK 24

Db 292 GETYQCRVTHPLPRALVRSSTK 314

##### RESULT 2

EHHU

Ig epsilon chain C region - human

C:Species: Homo sapiens (man)

C>Date: 31-Mar-1981 #sequence\_revision 13-Jun-1983 #text\_change 16-Jul-1999

C/Accession: A22771; A23195; P1214; A93491; A90824; A94418; B93933; S02438; A53116; C4

R:Flanagan, J.G.; Rabbitts, T.H.

EMBO J. 1, 655-660, 1982

A:Title: The sequence of a human immunoglobulin epsilon heavy chain constant region gen

A:Reference number: A22771; MUID:84236029; PMID:6234164

A/Accession: A22771

A:Molecule type: DNA

A:Residues: 1-428 <FLA>

A/Cross-references: GB:I00022; GB:J00227; GB:V00555; NID:G185035

R:Ueda, S.; Nakai, S.; Nishida, Y.; Hisajima, H.; Honjo, T.

EMBO J. 1, 1533-1544, 1982

A:Title: Long terminal repeat-like elements flank a human immunoglobulin epsilon pseudo

A:Reference number: A23195; MUID:84207910; PMID:6327276

A/Accession: A23195

A:Molecule type: DNA

A:Residues: 2-428 <UED>

A/Cross-references: GB:J00222; NID:G184755

R:Zhang, K.; Saxon, A.; Max, E.E.

J. Exp. Med. 176, 233-243, 1992  
A/Title: Two unusual forms of human immunoglobulin E encoded by alternative RNA splicing  
A/Reference number: PH1214; MUID:92308839; PMID:1613458  
A/Accession: PH1214  
A/Molecule type: DNA  
A/Residues: 320-428 <ZHA>  
A/Cross-references: EMBL:X63693; GB:S38668; NID:G32987  
R/Seno, M.; Kurokawa, T.; Ono, Y.; Onda, H.; Sasada, R.; Igarashi, K.; Kikuchi, M.; Sugii  
Nucleic Acids Res. 11, 719-726, 1983  
A/Title: Molecular cloning and nucleotide sequencing of human immunoglobulin epsilon chain  
A/Reference number: A93491; MUID:83168897; PMID:6300763  
A/Accession: A93491  
A/Molecule type: mRNA  
A/Residues: 1-428 <SEN>  
A/Cross-references: GB:I00022; GB:J00227; GB:V00555; NID:G185035  
R/Max, E.E.; Batey, J.; Ney, R.; Kirsch, I.R.; Leder, P.  
Cell 29, 691-699, 1982  
A/Title: Duplication and deletion in the human immunoglobulin epsilon genes.  
A/Reference number: A90824; MUID:83001945; PMID:6288268  
A/Accession: A90824  
A/Molecule type: DNA  
A/Residues: 1-358, 'L', 360-428 <MAX>  
A/Cross-references: GB:J00222; NID:G184755  
A/Note: This sequence difference may be due to polymorphism  
R/Bennich, H.H.; Johanson, S.G.O.; Von Bahr-Lindstrom, H.  
in Immediate Hypersensitivity: Modern Concepts and Developments, Bach, M.K., ed., pp.1-3  
A/Reference number: A94418  
A/Accession: A94418  
A/Molecule type: protein  
A/Residues: 'GAWTL', 6, 'X', 8-16, 'B', 45-52, 55-92, 95-97, 'B', 99-121, 'B', 123, 'L', 12  
A/Experimental source: myeloma protein Nd  
R/Kenten, J.H.; Molgaard, H.V.; Houghton, M.; Derbyshire, R.B.; Viney, J.; Bell, L.O.; G  
Proc. Natl. Acad. Sci. U.S.A. 79, 6661-6665, 1982  
A/Title: Cloning and sequence determination of the gene for the human immunoglobulin epsilon  
A/Reference number: A93933; MUID:83065234; PMID:6815656  
A/Accession: B93933  
A/Molecule type: mRNA  
A/Residues: 1-40; 68-114; 427-428 <KEN>  
A/Cross-references: GB:I00022; NID:G185035  
R/Ikeyama, S.  
FEBS Lett. 224, 306-310, 1987  
A/Title: Purification and characterization of a recombinant human IgE Fc-epsilon fragment  
A/Reference number: S02438; MUID:98083554; PMID:3121387  
A/Accession: S02438  
A/Status: nucleic acid sequence not shown  
A/Molecule type: mRNA  
A/Residues: 98-352 <IKE>  
R/Zhang, K.; Max, E.E.; Cheah, H.K.; Saxon, A.  
J. Biol. Chem. 269, 456-462, 1994  
A/Title: Complex alternative RNA splicing of epsilon-immunoglobulin transcripts produces  
A/Reference number: A53116; MUID:94103254; PMID:8276835  
A/Accession: A53116  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 320-428 <ZH2>  
A/Experimental source: myeloma U266-derived cell line AF-10  
R/Hellman, L.  
A/Note: sequence extracted from NCBI backbone (NCBIN:141701, NCBIP:141702)  
Eur. J. Immunol. 23, 159-167, 1993  
A/Title: Characterization of four novel epsilon chain mRNA and a comparative analysis of  
A/Reference number: A46536; MUID:93122085; PMID:8419166  
A/Accession: C46536  
A/Status: preliminary; not compared with conceptual translation  
A/Molecule type: mRNA  
A/Residues: 382-426 <HEI>  
A/Cross-references: GB:S55273; NID:G263166; PIDN:AAB24855.1; PID:G263167  
A/Experimental source: B cell myeloma U-266  
A/Note: sequence extracted from NCBI backbone (NCBIP:125297)  
A/Accession: D46536  
A/Status: preliminary; not compared with conceptual translation  
A/Molecule type: mRNA  
A/Residues: 382-391 <HE2>  
A/Cross-references: GB:S55276; NID:G263168; PIDN:AAB24858.1; PID:G263169

A/Experimental source: B cell myeloma U-266  
A/Note: sequence extracted from NCBI backbone (NCBIP:125299)  
A/Accession: A46536  
A/Status: preliminary; not compared with conceptual translation  
A/Molecule type: DNA  
A/Residues: 401-428 <HE3>  
A/Cross-references: GB:S53497; NID:G263162; PIDN:AAB24855.1; PID:G263163  
A/Experimental source: B cell myeloma U-266  
A/Note: sequence extracted from NCBI backbone (NCBIP:123483)  
C/Genetics:  
A/Gene: GDB:IGHE  
A/Cross-references: GDB:119335; OMIM:147180  
A/Map position: 14q32.33-14q32.33  
A/Introns: 1/1; 104/1; 211/1; 319/1  
C/Complex: An immunoglobulin heterotetramer subunit consists of two identical light (ka  
hain disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into 1  
C/Superfamily: immunoglobulin C region; immunoglobulin homology  
C/Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglob  
F:22-87/Domain: immunoglobulin homology <IM1>  
F:128-195/Domain: immunoglobulin homology <IM2>  
F:232-301/Domain: immunoglobulin homology <IM3>  
F:338-407/Domain: immunoglobulin homology <IM4>  
F:14/Disulfide bonds: interchain (to light chain) #status predicted  
F:15-105,29-85,135-193,239-299,345-405/Disulfide bonds: #status predicted  
F:21,49,99,146,252,275/Binding site: carbonyl site: (Asn) (covalent) #status experimental  
F:121,209/Disulfide bonds: interchain (to heavy chain) #status predicted  
Query Match 60.3%; Score 85; DB 1; Length 428;  
Best Local Similarity 65.2%; Pred. No. 0.00013;  
Matches 15; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
Qy 2 GETYSRVTHPLPKDVRISIAK 24  
||| ||||| :|||  
Db 294 GETYQCRVTHPLPRALMRSTTK 316  
RESULT 3  
EHRT  
Ig epsilon chain C region - rat  
C/Species: Rattus norvegicus (Norway rat)  
C/Date: 17-Dec-1982 #sequence revision 17-Dec-1982 #text\_change 16-Jul-1999  
C/Accession: A93442; A90937; A02143  
R/Hellman, L.; Pettersson, U.; Engstrom, A.; Karlsson, T.; Bennich, H.  
Nucleic Acids Res. 10, 6041-6049, 1982  
A/Title: Structure and evolution of the heavy chain from rat immunoglobulin E.  
A/Reference number: A93442; MUID:83064537; PMID:6292865  
A/Accession: A93442  
A/Molecule type: mRNA  
A/Residues: 1-429 <HEL>  
A/Experimental source: strain LOU/c/Wal, immunocytoma IR2  
R/Kindsvogel, W.R.; Reddy, E.P.; Moore, J.M.; Faust Jr., C.H.  
DNA 1, 335-343, 1982  
A/Title: A cloned cDNA probe for rat immunoglobulin epsilon heavy chain: construction,  
A/Reference number: A90937; MUID:83182019; PMID:6820340  
A/Contents: myeloma IR162  
A/Accession: A90937  
A/Molecule type: mRNA  
A/Residues: 'N', 169-307, 'L', 309-342 <KIN>  
C/Complex: An immunoglobulin heterotetramer subunit consists of two identical light (ka  
hain disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into 1  
C/Superfamily: immunoglobulin C region; immunoglobulin homology  
C/Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
F:19-80/Domain: immunoglobulin homology <IM1>  
F:118-186/Domain: immunoglobulin homology <IM2>  
F:223-291/Domain: immunoglobulin homology <IM3>  
F:327-398/Domain: immunoglobulin homology <IM4>  
F:46,99,170,240,265,369,419/Binding site: carbonyl site: (Asn) (covalent) #status predict  
Query Match 58.2%; Score 92; DB 1; Length 429;  
Best Local Similarity 69.6%; Pred. No. 0.00036;  
Matches 16; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
Qy 2 GETYSRVTHPLPKDVRISIAK 24



```

Db      284 GEGYQCRVDHPHPKPIVRSITK 306
|||||
RESULT 4
Igb chain C3 region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jan-2000
R:Shinkai, Y.; Nakauchi, H.; Honjo, T.; Okumura, K.
Immunogenetics 27, 288-292, 1988
A:Title: Mouse immunoglobulin allotypes: Multiple differences between the nucleic acid
A:Reference number: I54443; MUID:88152907; PMID:3346043
A:Accession: I68730
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-107 <RES>
A:Cross-references: GB:M22933; NID:g194464; PIDN:AAA37915.1; PID:g194469
C:Superfamily: immunoglobulin C region; immunoglobulin homology <IMM>
F:22-90/Domain: immunoglobulin homology <IMM>

Query Match      41.1%; Score 58; DB 2; Length 107;
Best Local Similarity 56.5%; Pred. No. 0.27;
Matches 13; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY      2 GETYSRVTHPLPKDIVRSIAK 24
|||||
Db      83 GYGQCIVDHPDFPKPIVRSITK 105
|||||

RESULT 5
Igb chain C3 region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jan-2000
R:Shinkai, Y.; Nakauchi, H.; Honjo, T.; Okumura, K.
Immunogenetics 27, 288-292, 1988
A:Title: Mouse immunoglobulin allotypes: Multiple differences between the nucleic acid
A:Reference number: I54443; MUID:88152907; PMID:3346043
A:Accession: I68726
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-107 <RES>
A:Cross-references: GB:M22930; NID:g194455; PIDN:AAA37911.1; PID:g194460
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:22-90/Domain: immunoglobulin homology <IMM>

Query Match      41.1%; Score 58; DB 2; Length 107;
Best Local Similarity 56.5%; Pred. No. 0.27;
Matches 13; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY      2 GETYSRVTHPLPKDIVRSIAK 24
|||||
Db      83 GYGQCIVDHPDFPKPIVRSITK 105
|||||

RESULT 6
EHMS
Igb epsilon chain C region (version 1) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 13-Jun-1993 #sequence_revision 13-Jun-1993 #text_change 16-Jul-1999
R:Liu, F.T.; Albrandt, K.; Sutcliffe, J.G.; Katz, D.H.
Proc. Natl. Acad. Sci. U.S.A. 79, 7852-7856, 1982
A:Title: Cloning and nucleotide sequence of mouse immunoglobulin epsilon chain cDNA.
A:Reference number: A02144; MUID:8311774; PMID:6818553
A:Accession: A02144
A:Molecule type: mRNA
A:Residues: 1-388 <LU>
A:Cross-references: GB:J00476; NID:g194875; PIDN:AAA38085.1; PID:g3879220
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 1:
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:1-44/Domain: immunoglobulin homology (fragment) <IMI>
F:81-149/Domain: immunoglobulin homology <IM2>
F:186-254/Domain: immunoglobulin homology <IM3>
F:290-361/Domain: immunoglobulin homology <IM4>
F:110,51,62,133,205,228,332,382/Binding site: carbohydrate (Aan) (covalent) #status pred.

Query Match      41.1%; Score 58; DB 1; Length 388;
Best Local Similarity 56.5%; Pred. No. 1;
Matches 13; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY      2 GETYSRVTHPLPKDIVRSIAK 24
|||||
Db      247 GYGQCIVDHPDFPKPIVRSITK 269
|||||

RESULT 7
S38864
Igb epsilon chain C region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 24-May-2001
C:Accession: S38864
R:Kipp, B.; Becker, W.; Schlaak, M.
submitted to the EMBL Data Library, November 1993
A:Description: Combination of a defined specificity and desired isotype by cloning of a
A:Reference number: S38864
A:Accession: S38864
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-548 <KIP>
A:Cross-references: EMBL:Z27397; NID:g416537; PIDN:CAA81788.1; PID:g940782
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:353-421/Domain: immunoglobulin homology <IMM>

Query Match      41.1%; Score 58; DB 2; Length 548;
Best Local Similarity 56.5%; Pred. No. 1.5;
Matches 13; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY      2 GETYSRVTHPLPKDIVRSIAK 24
|||||
Db      414 GYGQCIVDHPDFPKPIVRSITK 436
|||||

RESULT 8
S37909
hypothetical protein YKL084w - yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Dec-2002
C:Accession: S37909
R:Pohl, I.M.; Pohl, F.M.
submitted to the Protein Sequence Database, March 1994
A:Reference number: S37897
A:Accession: S37909
A:Molecule type: DNA
A:Residues: 1-116 <POH>
A:Cross-references: EMBL:Z28084; NID:g486122; PID:g486123; GSPDB:GN00011; MIPS:YKL084w
A:Experimental source: strain S288C
C:Genetics:
A:Gene: MIPS:YKL084w
A:Cross-references: SGD:S0001567
A:Map position: 11L
C:Superfamily: uncharacterized conserved protein

Query Match      38.3%; Score 54; DB 2; Length 116;
Best Local Similarity 52.6%; Pred. No. 1.1;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY      7 SRVTHPLPKDIVRSIAK 25
|||||
Db      15 SRVTHPLPKDIVRSIAK 33
|||||
```

## RESULT 9

T19571  
 hypothetical protein C29F7.1 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
 C:Accession: T19571

R:McMurray, A.  
 submitted to the EMBL Data Library, March 1997

A:Reference number: T19571

A:Accession: T19571

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-394 <WIL>

A:Cross-references: EMBL:Z92827; PIDN:CA807326.1; GSPDB:GN00028; CESP:C29F7.1

A:Experimental source: clone C29F7

C:Genetics:

A:Gene: CESP:C29F7.1

A:Map position: X

A:Introns: 32/3; 72/3; 117/1; 164/3; 193/3; 247/1; 335/3; 363/3

Query Match 38.3%; Score 54; DB 2; Length 394;

Best Local Similarity 60.0%; Pred. No. 4.1;

Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 11 HPNLPKDIVRSIAC 25

DB 62 HPNLPKDIVRSIAC 76

RESULT 10

S42659

acetyl-CoA carboxylase (EC 6.4.1.2) - maize (fragment)

C:Species: Zea mays (maize)

C>Date: 07-Sep-1994 #sequence\_revision 24-Oct-1998 #text\_change 17-Mar-1999

C:Accession: S42659; S34636

R:Ashton, A.R.; Jenkins, C.L.D.; Whitfield, P.R.

Plant Mol. Biol. 24, 35-49, 1994

A:Title: Molecular cloning of two different cDNAs for maize acetyl CoA carboxylase.

A:Reference number: S42659; MUID:94154242; PMID:7906562

A:Accession: S42659

A:Molecule type: mRNA

A:Residues: 1-1306 <ASH>

A:Cross-references: EMBL:Z24449

C:Function:

A:Description: catalyzes the ATP-dependent carboxylation of acetyl-CoA to malonyl-CoA

A:Pathway: fatty acid biosynthesis

C:Superfamily: human acetyl-CoA carboxylase; biotin carboxylase homology; lipoyl/biotin-

C:Keywords: ligase

Query Match 38.3%; Score 54; DB 2; Length 1306;

Best Local Similarity 55.6%; Pred. No. 14;

Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 ETVYSRVTHPLPKDIVR 20

DB 140 ETVYSRVTHPLPKDIVR 157

RESULT 11

T02921

acetyl-CoA carboxylase (EC 6.4.1.2) (clone A3) - maize (fragment)

C:Species: Zea mays (maize)

C>Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 21-Jul-2000

C:Accession: T02921

R:Ashton, A.R.; Jenkins, C.L.D.; Whitfield, P.R.

Plant Mol. Biol. 24, 35-49, 1994

A:Title: Molecular cloning of two different cDNAs for maize acetyl CoA carboxylase.

A:Reference number: S42659; MUID:94154242; PMID:7906562

A:Accession: T02921

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1625 <ASH>

A:Cross-references: EMBL:Z24449; NID:91272684; PIDN:CAA80822.1; PID:91272685  
 A:Experimental source: strain B73; leaf  
 C:Function:

A:Description: catalyzes the ATP-dependent carboxylation of acetyl-CoA to malonyl-CoA  
 C:Superfamily: human acetyl-CoA carboxylase; biotin carboxylase homology; lipoyl/biotin  
 C:Keywords: ligase  
 F:68-140/Domain: lipoyl/biotin-binding homology <LPB>

Query Match 38.3%; Score 54; DB 2; Length 1625;

Best Local Similarity 55.6%; Pred. No. 18;

Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 ETVYSRVTHPLPKDIVR 20

DB 459 ETVYSRVTHPLPKDIVR 476

RESULT 12

T06161

acetyl-CoA carboxylase (EC 6.4.1.2) - wheat

N:Alternate names: acetyl-coenzyme A carboxylase

C:Species: Triticum aestivum (common wheat)

C>Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 18-Aug-2000

C:Accession: T06161; T06162; S42660; S78600; S35959

R:Gornicki, P.; Paris, J.; Podkowinski, J.; Gill, B.; Haselkorn, R.

Proc. Natl. Acad. Sci. U.S.A. 94, 14179-14184, 1997

A:Title: Plastid localized acetyl-CoA carboxylase of bread wheat is encoded by a single

A:Reference number: Z15495; MUID:98054381; PMID:9391173

A:Accession: T06161

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-2311 <GOR1>

A:Cross-references: EMBL:AF029895; NID:92827149; PIDN:AAC39330.1; PID:92827150

A:Experimental source: cv. Hard Red Winter Tam 107, light-grown seedlings

A:Accession: T06162

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-10; F', 12-67, H', 69-502, T', 504-619 <GOR2>

A:Cross-references: EMBL:AF029897; NID:92827153; PIDN:AAC39332.1; PID:92827154

A:Experimental source: cv. Hard Red Winter Tam 107, light-grown seedling

R:Elborough, K.M.; Simon, J.W.; Swinhoe, R.; Ashton, A.R.; Slabas, A.R.

Plant Mol. Biol. 24, 21-34, 1994

A:Title: Studies on wheat acetyl CoA carboxylase and the cloning of a partial cDNA.

A:Reference number: S42660; MUID:94154234; PMID:7906561

A:Accession: S42660

A:Molecule type: mRNA

A:Residues: 'WRTYM', 11771-2026, 'K', 2028-2073, 'F', 2075-2108, 'V', 2110-2120, 'L', 2122, 'A', 21

A:Cross-references: EMBL:Z23038; NID:9396278; PIDN:CAA80573.1; PID:9396279

A:Accession: S78600

A:Molecule type: protein

A:Residues: 1847-1852, 'E', 1854-1863; 1947-1962, 'D', 1964; 2085-2091, 'HL', 2139-2155, 'I', 215

C:Genetics:

A:Gene: Acc-1

A:Introns: 69/1; 90/2; 209/3; 278/3; 332/3; 363/3; 395/3; 453/3; 485/1; 547/3; 564/3; 5

C:Function:

A:Description: catalyzes the ATP-dependent carboxylation of acetyl-CoA to malonyl-CoA

A:Pathway: fatty acid biosynthesis

C:Superfamily: human acetyl-CoA carboxylase; biotin carboxylase homology; lipoyl/biotin

C:Keywords: ligase

F:134-640/Domain: biotin carboxylase homology <BCH>

F:767-839/Domain: lipoyl/biotin-binding homology <LPB>

F:806/Binding site: biotin (lys) (covalent) #status predicted

Query Match 38.3%; Score 54; DB 2; Length 2311;

Best Local Similarity 55.6%; Pred. No. 26;

Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 ETVYSRVTHPLPKDIVR 20

DB 1159 ETVYSRVTHPLPKDIVR 1176

RESULT 13

T02235  
acetyl-CoA carboxylase (EC 6.4.1.2) - maize  
N/Alternate names: acetyl-coenzyme A carboxylase  
C/Species: Zea mays (maize)  
C/Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 21-Jan-2000  
C/Accession: T02235  
R/Eggl, M.A.; Lutz, S.M.; Somers, D.A.; Gengenbach, B.G.  
Plant Physiol. 108, 1299-1300, 1995  
A/Title: A maize acetyl-coenzyme A carboxylase cDNA sequence.  
A/Reference number: Z14632; MUID:95357420; PMID:7630949  
A/Accession: T02235  
A/Status: preliminary; translated from GB/EMBL/DDBJ  
A/Molecule type: mRNA  
A/Residues: 1-2325 <EG>  
A/Cross-references: EMBL:U19183; NID:gl045304; PIDN:AAA80214.1; PID:9854731  
A/Experimental source: strain A188; leaf  
C/Function:  
A/Description: catalyzes the ATP-dependent carboxylation of acetyl-CoA to malonyl-CoA  
A/Pathway: fatty acid biosynthesis  
C/Superfamily: human acetyl-CoA carboxylase; biotin carboxylase homology; lipoyl/biotin-  
C/Keywords: ligase  
F/334-640/Domain: biotin carboxylase homology <BC>  
F/767-839/Domain: lipoyl/biotin-binding homology <LP>  
Query Match 38.3%; Score 54; DB 2; Length 2325;  
Best Local Similarity 55.6%; Pred. No. 26;  
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
QY 3 ETYSRVTHPLPKDIVR 20  
DB 1158 ETYSRLYQPHLVKDSIQ 1175  
RESULT 14  
T02599  
hypotheical protein A2g14770 [imported] - Arabidopsis thaliana  
N/Alternate names: Hypotheical protein F26C24.9  
C/Species: Arabidopsis thaliana [mouse-ear cress]  
C/Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 16-Feb-2001  
C/Accession: T02599; C84521  
R/Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,  
submitted to the EMBL Data Library, June 1998  
A/Description: Arabidopsis thaliana chromosome II BAC F26C24 genomic sequence.  
A/Reference number: Z14680  
A/Accession: T02599  
A/Status: translated from GB/EMBL/DDBJ  
A/Molecule type: DNA  
A/Residues: 1-1756 <ROU>  
A/Cross-references: EMBL:AC004705; NID:g3252804; PID:g3252818  
A/Experimental source: cultivar Columbia  
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.B.; Umayam, L.; Tallon, L.;  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,  
Nature 402, 761-768, 1999  
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A/Reference number: A84420; MUID:20083487; PMID:10617197  
A/Accession: C84521  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-1756 <STO>  
A/Cross-references: GB:AE002093; NID:g3252818; PIDN:AAC24188.1; GSPDB:GN00139  
C/Genetics:  
A/Gene: At2g14770; F26C24.9  
A/Map position: 2  
A/Introns: 42/1; 351/3; 418/1; 617/3; 659/3; 690/3; 725/1; 758/3; 778/3; 812/3; 902/2; 9  
Query Match 36.9%; Score 52; DB 2; Length 1756;  
Best Local Similarity 37.5%; Pred. No. 38;  
Matches 9; Conservative 6; Mismatches 9; Indels 0; Gaps 0;  
QY 1 CGETYSRVTHPLPKDIVRSIAK 24  
DB 1337 CRNTLLNMTNPDPDLAKIAK 1360

## RESULT 15

## PFMS

alpha-fetoprotein precursor - mouse  
N/Alternate names: AFP; alpha-1-fetoprotein precursor; alpha-fetoglobulin precursor  
C/Species: Mus musculus (house mouse)  
C/Date: 01-Sep-1981 #sequence\_revision 28-May-1986 #text\_change 12-Jun-1998  
C/Accession: A93254; A92305; A03235  
R/Law, S.W.; Dugaiczky, A.  
Nature 291, 201-205, 1981  
A/Title: Homology between the primary structure of alpha-fetoprotein, deduced from a c  
A/Reference number: A93254; MUID:81197641; PMID:6164927  
A/Accession: A93254  
A/Molecule type: mRNA  
A/Residues: 1-597; E, 599-605 <LAW>  
A/Cross-references: GB:V00743  
R/Minghetti, P.P.; Law, S.W.; Dugaiczky, A.  
Mol. Biol. Evol. 2, 347-358, 1985  
A/Title: The rate of molecular evolution of alpha-fetoprotein approaches that of pseudo;  
A/Reference number: A93055; MUID:88216123; PMID:2452956  
A/Contents: annotation; revision to residue 598  
A/Note: residue 598, reported in reference A93254 as Glu, should have been reported as  
R/Gorin, M.B.; Cooper, D.L.; Eiferman, F.; van de Rijdt, P.; Tilghman, S.M.  
J. Biol. Chem. 256, 1954-1959, 1981  
A/Title: The evolution of alpha-fetoprotein and albumin: I. A comparison of the primary  
A/Reference number: A92305; MUID:81117287; PMID:6161929  
A/Accession: A92305  
A/Molecule type: mRNA  
A/Residues: 15-532; 'RAKL', 538-605 <GOR>  
A/Cross-references: GB:M16381  
A/Note: the beginning of the mature protein was placed at residue 21  
R/Eiferman, F.A.; Young, P.R.; Scott, R.W.; Tilghman, S.M.  
Nature 294, 713-718, 1981  
A/Title: Intragenic amplification and divergence in the mouse alpha-fetoprotein gene.  
A/Reference number: A93271; MUID:82103990; PMID:6172714  
A/Contents: annotation; exon-intron junctions  
C/Comment: Mouse AFP has two carbohydrate chains and was found to bind estrogens and co  
C/Genetics:  
A/Map position: 5  
A/Introns: 29/1; 46/2; 86/3; 157/2; 201/3; 234/2; 277/3; 349/2; 393/3; 426/2; 472/3; 54  
C/Superfamily: serum albumin; serum albumin repeat homology  
C/Keywords: embryo; fetus; globulin; glycoprotein; metal binding; plasma  
F/1-20/Domain: signal sequence #status predicted <SIG>  
F/21-605/Product: alpha-fetoprotein #status predicted <MPT>  
F/29-198/Domain: serum albumin repeat homology <SA1>  
F/217-390/Domain: serum albumin repeat homology <SA2>  
F/409-598/Domain: serum albumin repeat homology <SA3>  
F/22/Binding site: copper (His) #status predicted  
F/95-110,109-120,144-189,188-197,220-266,265-273,285-299,298-309,380-389,412-458,457-46  
F/247,498/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 36.5%; Score 51.5; DB 1; Length 605;  
Best Local Similarity 57.9%; Pred. No. 15;  
Matches 11; Conservative 3; Mismatches 4; Indels 1; Gaps 1;  
QY 6 YSRVTHPLPKDIVRSIAK 24  
DB 354 YSR-THPLPVSVILRIAK 371  
Search completed: February 26, 2004, 08:28:13  
Job time : 12 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 08:17:40 ; Search time 6.4 Seconds  
(without alignments)  
203.399 Million cell updates/sec

Title: US-09-701-623C-6  
Perfect score: 141  
Sequence: 1 CGETYSRVTHPLPKDIVERIAKC 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	85	60.3	428	1 EPC HUMAN	P01854 homo sapien
2	82	58.2	429	1 EPC RAT	P01855 rattus norv
3	58	41.1	421	1 EPC MOUSE	P06336 mus musculu
4	54	38.3	116	1 YK14 YEAST	P36078 saccharomyc
5	51.5	36.5	605	1 FETA MOUSE	P02772 mus musculu
6	51	36.2	106	1 Y116 ADE02	P03287 human adeno
7	50	35.5	518	1 GVIC CAEBR	Q8muk3 caenorhabdi
8	49.5	35.1	381	1 GVIC HALME	Q02228 halobacteri
9	49.5	35.1	382	1 GVIC HALN1	P24574 halobacteri
10	48	34.0	384	1 GBA1_PEA	O04279 pisum sativ
11	48	34.0	1161	1 POL_SFV1	P23074 simian foam
12	47.5	33.7	327	1 SYW_PORGI	Q7mt94 porphyromon
13	47	33.3	352	1 MODC SALT1	Q828a4 salmonella
14	47	33.3	352	1 MODC SALT2	Q828a6 salmonella
15	47	33.3	384	1 GBA1_LYCES	P26981 lycopersico
16	47	33.3	777	1 BARI_HUMAN	Q99728 homo sapien
17	46	32.6	264	1 H82D_RAT	P18211 rattus norv
18	46	32.6	607	1 UVRC_PSEFL	P32966 pseudomonas
19	45.5	32.3	270	1 KORB METJA	Q59577 methanococc
20	45.5	32.3	371	1 TRMU OCEBI	Q8cxc7 oceanobacil
21	45	31.9	242	1 VMA1_INGJU	P12446 influenza c
22	45	31.9	319	1 CDK4_XENLA	Q91727 xenopus lae
23	45	31.9	428	1 FX3A_HUMAN	Q9ukt7 homo sapien
24	45	31.9	428	1 FX3A_MOUSE	Q8c4v4 mus musculu
25	45	31.9	626	1 UVRC_RICPR	Q9zcx9 rickettsia
26	45	31.6	639	1 UVRC_RICCN	Q92ha7 rickettsia
27	44.5	31.6	1235	1 IRS1_RAT	P35570 rattus norv
28	44	31.2	161	1 COAD_STRA3	Q8e6f1 streptococc
29	44	31.2	285	1 GVC2_HALN1	Q9hbt0 halobacteri
30	44	31.2	285	1 GVC2_HALN2	Q48310 halobacteri
31	44	31.2	374	1 SPOF_HUMAN	Q43791 homo sapien
32	44	31.2	531	1 PUOR_BACSU	C02138 bacillus su
33	43.5	30.9	429	1 WR44_ARATH	Q9zuu0 arabidopsis

RESULT 1				
EPC_HUMAN				
ID	EPC_HUMAN	STANDARD;	PRT;	428 AA.
AC	P01854;			
DT	21-JUL-1986	(Rel. 01, Created)		
DT	21-JUL-1986	(Rel. 01, Last sequence update)		
DT	10-OCT-2003	(Rel. 42, Last annotation update)		
DE	Ig epsilon chain C region.			
GN	IGHF.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=83168897; PubMed=6300763;			
RA	Seno M., Kurokawa T., Ono Y., Onda H., Sasada R., Igarashi K.,			
RA	Kikuchi M., Sugino Y., Nishida Y., Honjo T.;			
RT	"Molecular cloning and nucleotide sequencing of human immunoglobulin			
RT	epsilon chain cDNA.";			
RL	Nucleic Acids Res. 11:719-726(1983).			
RN	[2]			
RP	SEQUENCE FROM N.A., AND VARIANT LEU-359.			
RX	MEDLINE=83001945; PubMed=6288268;			
RA	Max E.E., Battey J., Ney R., Kirsch I.R., Leder P.;			
RT	"Duplication and deletion in the human immunoglobulin epsilon genes.";			
RL	Cell 29:691-699(1982).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=842336029; PubMed=62341164;			
RA	Flanagan J.G., Rabbitts T.H.;			
RT	"The sequence of a human immunoglobulin epsilon heavy chain constant			
RT	region gene, and evidence for three non-allelic genes.";			
RL	EMBO J. 1:655-660(1982).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=84207910; PubMed=6327276;			
RA	Ueda S., Nakai S., Nishida Y., Hisajima H., Honjo T.;			
RT	"Long terminal repeat-like elements flank a human immunoglobulin			
RT	epsilon pseudogene that lacks introns.";			
RL	EMBO J. 1:1539-1544(1982).			
RN	[5]			
RP	PRELIMINARY SEQUENCE (MYELOMA PROTEIN ND).			
RA	Bennich H.H., Johanson S.G.O., von Bahr-Lindstrom H.;			
RA	(In) Bach M.K. (eds.);			
RL	Immediate hypersensitivity: modern concepts and developments, pp.1-36,			
RL	Marcel Dekker, New York (1978).			
RN	[6]			
RP	SEQUENCE OF 1-40; 68-114 AND 427-428 FROM N.A.			
RX	MEDLINE=83065234; PubMed=6815656;			
RA	Kenten J.H., Molgaard H.V., Houghton M., Derbyshire R.B., Viney J.,			
RA	Bell L.O., Gould H.J.;			
RT	"Cloning and sequence determination of the gene for the human			
RT	immunoglobulin epsilon chain expressed in a myeloma cell line.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 79:6661-6665(1982).			
RN	[7]			

P01874 canis famil  
P20439 drosophila  
Q93274 xenopus lae  
P08687 pisum sativ  
Q97K22 clostridium  
P13634 mus musculu  
Q58170 methanococc  
Q81J42 bacillus ce  
Q9B844 homo sapien  
O88886 mus musculu  
P04220 homo sapien  
Q04536 saccharomyc

## ALIGNMENTS

```
RP 3D-STRUCTURE MODELING.
RX MEDLINE=87089848; PubMed=3796618;
RA Padlan E.A., Davies D.R.;
RT "A model of the Fc of immunoglobulin E.";
RL Mol. Immunol. 23:1063-1075(1986).
CC -!- SIMILARITY: Contains 4 immunoglobulin-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L00022; RA859424.1; ALT_INIT.
CC PIR; A22771; EHHU.
CC PDB; 1IG5; 15-JUL-92.
CC PDB; 1FP5; 30-JAN-02.
CC PDB; 1G84; 16-MAY-01.
CC PDB; 1OOV; 18-SEP-02.
CC Genew; HGNC:5522; IGHE.
CC MIM; 147180; -.
CC GO; GO:0003823; F:antigen binding; NAS.
CC GO; GO:0006955; P:immune response; NAS.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003597; Ig-cl.
CC InterPro; IPR003006; Ig_MHC.
CC Pfam; PF00047; Ig; 4.
CC SMART; SM00407; IGc1; 4.
CC PROSITE; PS00835; IG_LIKE; 4.
CC PROSITE; PS00290; IG_MHC; 3.
CC Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW 3D-structure; Repeat.
FT NON_TER 1
FT DOMAIN 6
FT DOMAIN 112
FT DOMAIN 214
FT DOMAIN 324
FT DISULFID 14
FT DISULFID 15
FT DISULFID 29
FT DISULFID 121
FT DISULFID 135
FT DISULFID 209
FT DISULFID 239
FT DISULFID 345
FT CARBOHYD 21
FT CARBOHYD 49
FT CARBOHYD 99
FT CARBOHYD 146
FT CARBOHYD 252
FT CARBOHYD 275
FT VARIANT 359
FT STRAND 110
FT STRAND 113
FT STRAND 122
FT TURN 125
FT STRAND 130
FT STRAND 146
FT STRAND 155
FT STRAND 159
FT STRAND 165
FT TURN 168
FT STRAND 172
FT STRAND 182
FT TURN 187
FT STRAND 192
FT TURN 198
FT TURN 201
FT STRAND 205
FT STRAND 218
FT IG-LIKE 1.
FT IG-LIKE 2.
FT IG-LIKE 3.
FT IG-LIKE 4.
FT INTERCHAIN (WITH A LIGHT CHAIN).
FT INTERCHAIN (WITH A HEAVY CHAIN).
FT INTERCHAIN (WITH A HEAVY CHAIN).
FT N-LINKED (GLCNAC. . .).
FT N-LINKED (GLCNAC. . .).
FT N-LINKED (GLCNAC. . .).
FT N-LINKED (GLCNAC. . .).
FT N-LINKED (GLCNAC. . .).
FT W -> L (POSSIBLE POLYMORPHISM).
/FTid=VAR_003885.
FT 110
FT 117
FT 122
FT 125
FT 130
FT 141
FT 150
FT 155
FT 161
FT 166
FT 169
FT 181
FT 186
FT 188
FT 196
FT 199
FT 202
FT 207
FT 222
```

```
FT TURN 226
FT TURN 231
FT STRAND 236
FT STRAND 242
FT STRAND 252
FT STRAND 255
FT TURN 257
FT STRAND 258
FT STRAND 260
FT TURN 266
FT STRAND 271
FT TURN 274
FT STRAND 276
FT STRAND 277
FT STRAND 285
FT HELIX 288
FT TURN 293
FT TURN 294
FT STRAND 297
FT TURN 302
FT TURN 304
FT STRAND 310
FT STRAND 314
FT STRAND 322
FT STRAND 325
FT STRAND 334
FT STRAND 337
FT STRAND 340
FT STRAND 348
FT STRAND 350
FT STRAND 351
FT TURN 356
FT TURN 366
FT STRAND 374
FT STRAND 377
FT STRAND 380
FT TURN 381
FT STRAND 384
FT STRAND 385
FT STRAND 387
FT STRAND 393
FT TURN 394
FT TURN 398
FT TURN 399
FT STRAND 404
FT STRAND 408
FT STRAND 418
FT STRAND 421
SQ SEQUENCE 428 AA; 47019 MW; 25C4CA072AA558A0 CRC64;
```

Query Match 60.3%; Score 85; DB 1; Length 428;

Best Local Similarity 65.2%; Pred. No. 4.4e-05; Matches 15; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 GETYSRVTHPLPKDIVRSIAK 24  
|||||  
DB 294 GETYQCRVTHPLPRALMRSTTK 316  
|||||

#### RESULT 2

```
EPC_RAT
ID EPC_RAT STANDARD; PRT; 429 AA.
AC P01855;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig epsilon chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (IMMUNOCYTOMA IR2).
RC STRAIN=LOU/C/WSL;
RX MEDLINE=83064537; PubMed=6292865;
RA Hellman L., Petterson U., Engstroem A., Karlsson T., Bennich H.;
RT "Structure and evolution of the heavy chain from rat immunoglobulin
RT E."
RL Nucleic Acids Res. 10:6041-6049(1982).
RN [2]
RP SEQUENCE OF 168-342 FROM N.A. (MYELOMA IR162).
RX MEDLINE=83182019; PubMed=6620340;
RA Kindsvogel W.R., Reddy E.P., Moore J.M., Faust C.H. Jr.;
RT "A cloned cDNA probe for rat immunoglobulin epsilon heavy chain:
RT construction, identification, and DNA sequence."
RL DNA 1:335-343(1982).
RN [3]
RP SEQUENCE OF 205-306 FROM N.A.
RX MEDLINE=82174576; PubMed=6603238;
```

```

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.ebi.ac.uk/ebis-sib.ch/announcement or send an email to licensese@ebi.ac.uk).
-----
EMBL; X01857; CAA25977.1; -
EMBL; X01857; CAA25978.1; -
PIR; A02144; EHMS.
PIR; A02145; EHMS.
HSP; P01854; LIIE.
InterPro; IPR007110; Ig-Like.
InterPro; IPR003597; Ig cl.
InterPro; IPR003006; Ig_MHC.
Pfam; PF00047; Ig; 4.
SMART; SM00407; IGcl; 2.
PROSITE; PS00835; IG_LIKE; 4.
PROSITE; PS00290; IG_MHC; 3.
Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
NON_TER 1 1
DOMAIN 1 90 CH1.
FT DOMAIN 91 197 CH2.
FT DOMAIN 198 304 CH3.
FT DOMAIN 305 421 CH4.
FT DISULFID 23 75 BY SIMILARITY.
FT DISULFID 121 180 BY SIMILARITY.
FT DISULFID 226 285 BY SIMILARITY.
FT DISULFID 330 392 BY SIMILARITY.
FT CARBOHYD 43 43 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 95 95 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 261 261 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 365 365 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 415 415 N-LINKED (GLCNAC. .) (POTENTIAL).
FT SEQUENCE 421 AA; 47320 MW; 8F909E1F30A06B47 CRC64;
Query Match 41.1%; Score 58; DB 1; Length 421;
Best Local Similarity 56.5%; Pred. No. 0.43;
Matches 13; Conservative 0; Mismatches 0; Indels 10; Gaps 0;
QY 2 GETYSRVTHPLPKOIVRSIAK 24
| | | | | | | | | | | | | |
Db 280 GYGVCIVDPDFPFKPVRISIK 302
-----
RESULT 4
YK14_YEAST
ID YK14_YEAST STANDARD; PRT; 116 AA.
AC P36078;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Hypothetical 13.6 kDa protein in MDH1-VMA5 intergenic region.
GN YKL084W
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
NCBI_TaxId=4932;
RN [1]
SEQUENCE FROM N.A.
RP Pohl T.M., Pohl F.M.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

```

FT	SIGNAL	1	18	PROBABLE.
FT	CHAIN	19	605	ALPHA-FETOPROTEIN
FT	CHAIN	19	605	ALBUMIN 1.
FT	DOMAIN	201	201	ALBUMIN 1.
FT	DOMAIN	393	393	ALBUMIN 2.
FT	DOMAIN	400	591	ALBUMIN 3.
FT	DISULFID	95	110	
FT	DISULFID	109	120	
FT	DISULFID	144	189	
FT	DISULFID	188	197	
FT	DISULFID	220	266	

```

FT DISULFID 265 273
FT DISULFID 285 299
FT DISULFID 298 309
FT DISULFID 380 389
FT DISULFID 412 458
FT DISULFID 457 468
FT DISULFID 481 497
FT DISULFID 496 507
FT DISULFID 534 579
FT DISULFID 578 587
FT CARBOHYD 247 247
FT CARBOHYD 325 325
FT CARBOHYD 498 498
FT CONFLICT 533 536
FT SEQUENCE 605 AA; 67337 MW; CE09E9F50D74619A CRC64;

Query Match 36.5%; Score 51.5; DB 1; Length 605;
Best Local Similarity 57.9%; Pred. No. 5.9;
Matches 11; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

QY 6 YSRVTHPLPKDIVRSIAK 24
DB 354 YSR-THPNLPVSVILRIAK 371

RESULT 6
Y116 ADE02 STANDARD; PRT; 106 AA.
AC P03267;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-APR-1990 (Rel. 14, Last annotation update)
DE Hypothetical 11.6 kDa early protein.
OS Human adenovirus type 2.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=10515;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83056843; PubMed=7142161;
RA Gingras T.R., Slatky D., Gellinas R.E., Bing-Dong J., Yen C.E.,
RA Kelly M.M., Bullock P.A., Parsons B.L., O'Neill K.E., Roberts R.J.;
RT "Nucleotide sequences from the adenovirus-2 genome.";
RL J. Biol. Chem. 257:13475-13491(1982).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=83056844; PubMed=7142162;
RA Alestrom P., Akusjaervi G., Pettersson M., Pettersson U.;
RT "DNA sequence analysis of the region encoding the terminal protein
and the hypothetical N-gene product of adenovirus type 2.";
RL J. Biol. Chem. 257:13492-13498(1982).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL; AF520617; AAM78548.1;
DR GO; GO:0005634; C:nucleus; ISS.
DR GO; GO:0016538; P:cyclin-dependent protein kinase, intrinsic . . .; ISS.
DR GO; GO:0000082; P:G1/S transition of mitotic cell cycle; ISS.
DR GO; GO:0007088; P:regulation of mitosis; ISS.
DR InterPro; IPR006670; Cyclin.
DR InterPro; IPR006671; Cyclin_N.
DR Pfam; PF00134; cyclin; 1.
DR SMART; SM00385; CYCLIN; 1.
DR Cyclin; Cell cycle; Cell division; Nuclear protein.
KW SEQUENCE 518 AA; 59469 MW; A0849438E8891D3 CRC64;
SQ SEQUENCE 518 AA; 59469 MW; A0849438E8891D3 CRC64;
Query Match 35.5%; Score 50; DB 1; Length 518;
Best Local Similarity 50.0%; Pred. No. 8.3;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 6 YSRVTHPLPKDIVRSIA 23
DB 464 YDRFTQEHVPKDIIPDYA 481

RESULT 8
GVPC HALME STANDARD; PRT; 381 AA.
AC Q02228;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Gas vesicle protein C.
GN GVPC.
OS Halobacterium mediterranei (Haloflex mediterranei).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Haloflex.
OX NCBI_TaxID=2252;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 1411;

```



```

RX MEDLINE=93021102; PubMed=1404376;
RA Englert C., Krueger K., Offner S., Pfeifer F.;
RT "Three different but related gene clusters encoding gas vesicles in
RT halophilic archaea.";
RL J. Mol. Biol. 227:586-592(1992).
CC -!- FUNCTION: May confer stability to the gas vesicle membranes. Gas
CC vesicles are small, hollow, gas filled protein structures that are
CC found in several microbial planktonic microorganisms. They allow
CC the positioning of the organism at the favorable depth for growth.
CC -!- SUBCELLULAR LOCATION: Binds to the external surface of the gas
CC vesicle membrane.
CC -!- SIMILARITY: VERY DISTANT, TO GAS VESICLE PROTEINS TYPE C OF
CC CYANOBACTERIA.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X64701; CAA45944.1; -
DR PIR; S28115; S28115.
DR InterPro; IPR008639; Halo GVPC.
DR Pfam; PF05465; Halo GVPC; 1.
KW Gas vesicle; Repeat.
FT DOMAIN 22 274 7 X APPROXIMATE TANDEM REPEATS.
FT REPEAT 22 59 1.
FT REPEAT 22 84 2.
FT REPEAT 60 84 2.
FT REPEAT 85 122 3.
FT REPEAT 123 160 4.
FT REPEAT 161 192 5.
FT REPEAT 193 232 6.
FT REPEAT 233 274 7.
FT REPEAT 246 335 ASP/GLU-RICH (ACIDIC).
SQ SEQUENCE 381 AA; 42653 MW; 9FB48199D0305921 CRC64;
Query Match 35.1%; Score 49.5; DB 1; Length 381;
Best Local Similarity 64.3%; Pred. No. 7.1;
Matches 9; Conservative 1; Mismatches 3; Indels 1; Gaps 1;
Qy 1 CGEYYSRVTHPHL 14
Dy 343 CGE-YYQAITEPHL 355

```

## RESULT 9

```

GVCI_HALNI STANDARD; PRT; 382 AA.
AC P24574; Q9H117;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Gas vesicle protein C 1.
GN (GVPC1 OR GVPC OR VNG5032G) AND (GVPC2 OR VNG6031G).
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081), and
OS Halobacterium sp. (strain NRC-817).
OG Plasmid pNRC100, Plasmid pNRC200, and Plasmid pH1.
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
CX NCBI_TaxID=64091; 148370;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRC-1; PLASMID=pNRC100;
RX MEDLINE=91323716; PubMed=1864501;
RA Jones J.G., Young D.C., Dassarma S.;
RT "Structure and organization of the gas vesicle gene cluster on the
RT Halobacterium halobium plasmid pNRC100.";
RL Gene 102:117-122(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NRC-1; PLASMID=pNRC100;

```

```

RX MEDLINE=99063795; PubMed=9847077;
RA Ng W.V., Ciuto S.A., Smith T.M., Bungarner R.E., Baskin D., Faust J.,
RA Hall B., Loretz C., Sato J., Slagel J., Hood L., Dassarma S.;
RT "Snapshot of a large dynamic replicon in a halophilic archaeon:
RT megaplasmid or minichromosome?";
RL Genome Res. 8:1131-1141(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=NRC-1; PLASMID=pNRC200;
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Leaky S.R., Balliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Madocks D.G., Jablonki P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenberger T.A., Peck R.F., Fohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=NRC-817; PLASMID=pH1.
RX MEDLINE=92065812; PubMed=1956294;
RA Horne M., Englert C., Wimmer C., Pfeifer F.;
RT "A DNA region of 9 kbp contains all genes necessary for gas vesicle
RT synthesis in halophilic archaeobacteria.";
RL Mol. Microbiol. 5:1159-1174(1991).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=NRC-817; PLASMID=pH1.
RX MEDLINE=93021102; PubMed=1404376;
RA Englert C., Krueger K., Offner S., Pfeifer F.;
RT "Three different but related gene clusters encoding gas vesicles in
RT halophilic archaea.";
RL J. Mol. Biol. 227:586-592(1992).
CC -!- FUNCTION: May confer stability to the gas vesicle membranes. Gas
CC vesicles are small, hollow, gas filled protein structures that are
CC found in several microbial planktonic microorganisms. They allow
CC the positioning of the organism at the favorable depth for growth.
CC -!- SUBCELLULAR LOCATION: Binds to the external surface of the gas
CC vesicle membrane.
CC -!- SIMILARITY: VERY DISTANT, TO GAS VESICLE PROTEINS TYPE C OF
CC CYANOBACTERIA.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M58557; AAA98197.1; -
DR EMBL; AF016485; AAC82810.1; -
DR EMBL; AE005142; AAG20727.1; -
DR EMBL; X57161; CAA40451.1; -
DR EMBL; X64729; CAA45981.1; -
DR PIR; T08243; T08243.
DR InterPro; IPR008639; Halo GVPC.
DR Pfam; PF05465; Halo GVPC; 1.
KW Gas vesicle; Plasmid; Repeat; Complete proteome.
FT DOMAIN 22 284 7 X APPROXIMATE TANDEM REPEATS.
FT REPEAT 22 60 1.
FT REPEAT 61 92 2.
FT REPEAT 93 130 3.
FT REPEAT 131 168 4.
FT REPEAT 169 200 5.
FT REPEAT 201 240 6.
FT REPEAT 241 284 7.
FT DOMAIN 254 336 ASP/GLU-RICH (ACIDIC).
SQ SEQUENCE 382 AA; 42391 MW; 171DBEB4C0364F46 CRC64;

```

```

Query Match      35.1%; Score 49.5; DB 1; Length 382;
Best Local Similarity 64.3%; Pred.No.7.1.;
Matches          9; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY      1 CGETYSRVTHPHL 14
DB      344 CGE-YQAITEPAL 356
        |||||
RESULT 10
ID      ID_GBA2_PEA         STANDARD; PRT; 384 AA.
GBA2_PEA
AC      O04279;
DT      15-DEC-1998 (Rel. 37, Created)
DT      15-DEC-1998 (Rel. 37, Last sequence update)
DE      15-MAR-2004 (Rel. 43, Last annotation update)
DE      Guanine nucleotide-binding protein alpha-2 subunit (GP-alpha-2).
GN      GPA2 OR GA2.
OS      Pisum sativum (Garden pea).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC      eurosoids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
NCBI_TaxId=3889;
[1]
RN      SEQUENCE FROM N.A.
RC      STRAIN=cv Alaska; TISSUE=Apical bud;
RX      MEDLINE=99406511; PubMed=10476071;
RA      Marsh J.F. III, Kaufman L.S.;
RT      "Cloning and characterisation of PGA1 and PGA2: two G protein
RT      alpha-subunits from pea that promote growth in the yeast
RT      Saccharomyces cerevisiae.";
RL      Plant J. 19:237-247(1999).
CC      -I- FUNCTION: Guanine nucleotide-binding proteins (G proteins) are
CC      involved as modulators or transducers in various transmembrane
CC      signalling systems.
CC      -I- SUBUNIT: G proteins are composed of 3 units; alpha, beta and
CC      gamma. The alpha chain contains the guanine nucleotide binding
CC      site.
CC      -I- SIMILARITY: Belongs to the G-alpha family. Subfamily 2 (G(s)).
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
CC      EMBL; U97044; AAB57826.1; -.
CC      DR      PIR; T06814; T06814.
CC      DR      HSPSP; P10824; IAS3.
CC      DR      InterPro; IPR001019; Gprotein_alpha.
CC      DR      Pfam; PF00503; G-alpha_1.
CC      DR      PRINTS; PR00318; GPOTEINA.
CC      DR      ProDom; PD000281; Gprotein_alpha; 1.
CC      DR      SMART; SMO0275; G_alpha; 1.
CC      DR      GTP-binding; Transducer; ADP-ribosylation.
CC      FT      NP_BIND 46 53 GTP (BY SIMILARITY).
CC      FT      NP_BIND 219 223 GTP (BY SIMILARITY).
CC      FT      NP_BIND 288 291 GTP (BY SIMILARITY).
CC      FT      MOD_RES 191 191 ADP-RIBOSYL[1] (BY ACTION OF CTX).
CC      SQ      SEQUENCE 384 AA; 44669 MW; 6B0963393FD20B44 CRC64;

Query Match      34.0%; Score 48; DB 1; Length 384;
Best Local Similarity 50.0%; Pred.No.12;
Matches          8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```

```

POL_SFV1
ID POL_STANDARD; PRT; 1161 AA.
AC F23074;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE POL_P003protein [Contains: Protease (EC 3.4.23.-); Reverse
DE transcriptase (EC 2.7.7.49); Endonuclease].
DE POL.
OS Simian foamy virus (type 1) (SFV-1).
OS Viruses; Retroid viruses, Retroviridae; Spumavirus.
OX NCBI_TaxID=11964;
RN [1]
RN SEQUENCE FROM N.A.
RN RX MEDLINE=91276270; PubMed=1647358;
RA Kupiec J.-U., Kay A., Hayat M., Ravier R., Peries J., Galibert F.;
RA "Sequence analysis of the simian foamy virus type 1 genome.";
RT Gene 101:185-194(1991).
RN [2]
RN SEQUENCE OF 1-970 FROM N.A.
RN RX MEDLINE=91361538; PubMed=1653483;
RA Mergia A., Luciw P.A.;
RA "Replication and regulation of primate foamy viruses.";
RT Virology 184:475-482(1991).
RN [3]
RN SEQUENCE OF 969-1161 FROM N.A.
RN RX MEDLINE=90080148; PubMed=2152825;
RA Mergia A., Shaw K.E.S., Lackner J.R., Luciw P.A.;
RA "Relationship of the env genes and the endonuclease domain of the pol
RT genes of simian foamy virus type 1 and human foamy virus.";
RT J. Virol. 64:406-410(1990).
CC CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + [DNA] (N)
CC CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A9.
CC CC
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC ENBL; X54482; -; NOT_ANNOTATED_CDS.
DR DR
DR ENBL; X58484; CAA41394.1; -;
DR DR
DR ENBL; M33561; AAA47793.1; -;
DR DR
DR PIR; A33562; A33562.
DR DR
DR HSSP; P03355; LMML.
DR DR
DR MEROPS; A09.001; -.
DR DR
DR InterPro; IPR009007; Pept A acid.
DR DR
DR InterPro; IPR001641; Peptidase_A9.
DR DR
DR InterPro; IPR002156; RNaseH.
DR DR
DR InterPro; IPR001584; Rve.
DR DR
DR InterPro; IPR000477; RVTse.
DR DR
DR Pfam; PF00075; rnaseH; 1.
DR DR
DR Pfam; PF00665; rve; 1.
DR DR
DR Pfam; PF00078; rvt; 1.
DR DR
DR Pfam; PF03539; Spuma_A9ptase; 1.
DR DR
DR PRINTS; PR00920; SPUMVIRPTASE.
KW Transferase; RNA-directed DNA polymerase; Hydrolase; Nuclease;
KW Endonuclease; Aspartyl protease; Polyprotein.
KW ACT_SITE 36 36
KW FT CONFLICT 236 236
KW FT CONFLICT 910 910
KW FT CONFLICT 950 950
KW SEQUENCE 1161 AA; 131717 MW; 08D09D1DGBB4C14C CRC64;
SO

```

```
Query Match      34.0%; Score 48; DB 1; Length 1161;
Best Local Similarity 47.4%; Pred. No. 38;
Matches          9; Conservative      5; Mismatches 5; Indels 0; Caps 0;
```

## RESULT 11



```

RESULT 14.
MODC_SALTY
ID MODC_SALTY STANDARD; PRT; 352 AA.
AC Q8ZQ6;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Molybdenum import ATP-binding protein modC (BC 3.6.3.29).
GN MOC OR STM783.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2."
RL Nature 413:852-856 (2001).
CC -!- FUNCTION: Part of the ABC transporter complex modABC (TC 3.A.1.8)
CC involved in molybdenum import. Responsible for energy coupling to
CC the transport system (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + H2O + molybdate(Out) = ADP + phosphate
CC + molybdate(In).
CC -!- SUBUNIT: The complex is composed of two ATP-binding proteins
CC (modC), two transmembrane proteins (modB) and a solute-binding
CC protein (modA) (Probable).
CC -!- SUBCELLULAR LOCATION: Inner membrane-associated (By similarity).
CC -!- SIMILARITY: Belongs to the ABC transporter family. ModC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE008732; AAL19721.1; -.
CC SryGene; SG2277; modC.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR003439; ABC_Transporter.
CC InterPro; IPR004606; Mop.
CC InterPro; IPR008995; MOP like.
CC InterPro; IPR005116; TOBE.
CC Pfam; PF00005; ABC_tran; 1.
CC Pfam; PF03459; TOBE; 1.
CC SMART; SM00382; AAA; 1.
CC TIGRFAMs; TIGR00638; Mop; 1.
CC PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
CC PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
CC Hydrolase; Transport; Molybdenum; Membrane; Inner membrane;
CC ATP-binding; Complete proteome.
CC NP_BIND 31 38 ATP (By similarity).
SQ SEQUENCE 352 AA; 39055 MW; 65137E645711A85 CRC64;

Query Match 33.3%; Score 47; DB 1; Length 352;
Best Local Similarity 50.0%; Pred. No. 15;
Matches 11; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

OY 3 ETVYSRVTHPLPKDIVRSIAK 24
DB 215 EVGSSVWHPWLPKQOSSILK 236

RESULT 15
GBAI_LYCES
ID GBAI_LYCES STANDARD; PRT; 384 AA.
AC P26981;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Guanine nucleotide-binding protein alpha-1 subunit (GP-alpha-1).
GN GBAI OR GAI.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. VF36; TISSUE=Pistil;
RX MEDLINE=92084110; PubMed=1748292;
RA Ma H., Yancovsky M.F., Huang H.;
RT "Isolation and sequence analysis of TGA1 cDNAs encoding a tomato G
RT protein alpha subunit";
RL Gene 107:189-195 (1991).
CC -!- FUNCTION: Guanine nucleotide-binding proteins (G proteins) are
CC involved as modulators or transducers in various transmembrane
CC signaling systems.
CC -!- SUBUNIT: G proteins are composed of 3 units; alpha, beta and
CC gamma. The alpha chain contains the guanine nucleotide binding
CC site.
CC -!- SIMILARITY: Belongs to the G-alpha family. Subfamily 2 (G(s)).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M74419; AAA34167.1; -.
CC PIR; JH0514; RGT00A.
CC HSP; P10924; LAG3.
CC InterPro; IPR001019; Gprotein_alpha.
CC Pfam; PF00503; G-alpha; 1.
CC PRINTS; PR00318; GPROTEINA.
CC ProDom; PD000281; Gprotein_alpha; 1.
CC SMART; SM00275; G_alpha; 1.
CC GTP-binding; Transducer; ADP-ribosylation.
CC NP_BIND 46 53 GTP (By similarity).
CC NP_BIND 219 223 GTP (By similarity).
CC NP_BIND 288 291 GTP (By similarity).
CC MOD_RES 191 191 ADP-RIBOSYL[1] (BY ACTION OF CTX).
SQ SEQUENCE 384 AA; 44903 MW; 5C95ED74EFC01A0A CRC64;

Query Match 33.3%; Score 47; DB 1; Length 384;
Best Local Similarity 53.3%; Pred. No. 17;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 8 RVTHPLPKDIVRSI 22
DB 126 RLDVPHLTKDLVQDI 140

Search completed: February 26, 2004, 08:24:20
Job time : 8.4 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 08:18:11 ; Search time 31.4 Seconds  
(without alignments)  
251.209 Million cell updates/sec

Title: US-09-701-623C-6

Perfect score: 141  
Sequence: 1 CGETYSRVTHPLPKDIVRSIAKC 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

#### Database :

SPTREMBL 25:\*

- 1: sp\_archea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query	Score	Match	Length	DB ID	Description
1	57	40.4	1093	9	Q8SD16	Q8SD16 pseudomonas
2	54	38.3	394	5	O17620	O17620 caenorhabdi
3	54	38.3	1078	10	Q94FR6	Q94fr6 zea mays (m
4	54	38.3	1079	10	Q94FR7	Q94fr7 zea mays (m
5	54	38.3	1625	10	Q08367	Q08367 zea mays (m
6	54	38.3	2311	10	Q48959	Q48959 triticum ae
7	54	38.3	2324	10	Q7XVR4	Q7XVR4 zea mays (m
8	54	38.3	2325	10	Q41743	Q41743 zea mays (m
9	54	38.3	2325	10	Q7XVR3	Q7XVR3 zea mays (m
10	53	37.6	532	12	Q8JNC8	Q8Jnc8 macropodid
11	52	36.9	1756	10	O80975	O80975 arabidopsis
12	51.5	36.5	605	11	Q8BK65	Q8bk65 mus musculu
13	51.5	36.5	605	11	Q8BK56	Q8bk56 mus musculu
14	51.5	36.5	723	10	Q7XR27	Q7xr27 oryza sativ
15	51	36.2	195	17	Q9VIA1	Q9vial pyrococcus
16	51	36.2	416	2	Q9X4C5	Q9x4c5 escherichia

17	50	35.5	119	10	Q8CWM0	Q8cww0 arabidopsis
18	50	35.5	539	10	Q9LZX0	Q9lzx0 arabidopsis
19	49	34.8	405	11	Q91Z67	Q91z67 mus musculu
20	49	34.8	684	13	Q90544	Q90544 ginglymosto
21	49	34.8	1078	10	Q94FR5	Q94fr5 zea mays (m
22	49	34.8	1095	4	O75044	O75044 homo sapien
23	49	34.8	1695	10	Q43248	Q43248 zea mays (m
24	48	34.0	141	15	O87052	O87052 simian foam
25	48	34.0	152	12	O8JRY8	O8jry8 phthorimaea
26	48	34.0	367	16	Q97IP8	Q97ip8 clostridium
27	48	34.0	384	10	Q8LJQ8	Q8ljq8 pisum sativ
28	48	34.0	384	10	Q8LJQ8	Q8ljq8 streptomyce
29	48	34.0	387	16	Q82B97	Q82b97 streptomyce
30	48	34.0	388	16	Q9KY27	Q9ky27 drosophila
31	48	34.0	460	5	Q9VW78	Q9vw78 plasmodium
32	48	34.0	2240	5	O8IJ38	O8ij38 trichinella
33	47.5	33.7	421	5	Q9BIC9	Q9bic9 rhodopirell
34	47	33.3	81	16	Q7ULT4	Q7ult4 pyrococcus
35	47	33.3	195	17	Q8U0E1	Q8u0e1 arabidopsis
36	47	33.3	320	10	Q9LQ87	Q9lq87 arabidopsis
37	47	33.3	366	10	Q8W3P1	Q8w3p1 nicotiana t
38	47	33.3	366	10	Q8W3P2	Q8w3p2 nicotiana s
39	47	33.3	366	10	Q8W3P3	Q8w3p3 nicotiana t
40	47	33.3	372	10	Q40404	Q40404 nicotiana p
41	47	33.3	384	10	Q8W3P4	Q8w3p4 nicotiana t
42	47	33.3	384	10	Q9FV62	Q9fv62 nicotiana t
43	47	33.3	415	16	Q893F1	Q893f1 clostridium
44	47	33.3	427	10	Q8GYW7	Q8gyw7 arabidopsis
45	47	33.3	518	13	Q9W6R2	Q9w6r2 fugu rubrip

#### ALIGNMENTS

#### RESULT 1

Q8SD16 PRELIMINARY; PRT; 1093 AA.

AC Q8SD16; 01-JUN-2002 (TRENBLrel. 21, Created)

DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)

DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)

DE PHIKZ146.

OS Pseudomonas phage phiKZ.

OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.

OX NCBI\_TaxID=169683;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21914557; PubMed=11916376;

RA Mesyanzhinov V.V., Robben J., Grymonprez B., Kostyuchenko V.A.,

RA Bourkaltseva M.V., Sykilinda N.N., Krylov V.N., Volckaert G.;

RA Bourkaltseva M.V., Sykilinda N.N., Krylov V.N., Volckaert G.;

RT "The genome of bacteriophage phiKZ of Pseudomonas aeruginosa.";

RL J. Mol. Biol. 317:1-19(2002).

RN [2]

RP SEQUENCE FROM N.A.

RA Mesyanzhinov V.V., Robben J., Grymonprez B., Kostyuchenko V.A.;

RA Bourkaltseva M.V., Sykilinda N.N., Krylov V.V., Volckaert G.;

RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF395011; AAL63047.1; -

DR InterPro; IPR008160; Collagen.

DR Pfam; PF01391; Collagen; 2.

SQ SEQUENCE 1093 AA; 115835 MW; D755B6C6SD910DF7 CRC64;

Query Match 40.4%; Score 57; DB 9; Length 1093;

Best Local Similarity 39.1%; Pred. No. 6.6;

Matches 9; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 2 GETYSRVTHPLPKDIVRSIAK 24

DB 843 GKTEWATTFEIPSDVVRDTAQ 865

#### RESULT 2

O17620

```

ID O17620 PRELIMINARY; PRT; 394 AA.
AC O17620;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE C29F7.1 protein.
GN C29F7.1
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA McMurray A.A.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018 (1998).
DR EMBL; Z92827; CAB07326.1; -.
DR PIR; T19571; T19571.
DR WormPep; C29F7.1; CE08442.
DR InterPro; IPR004119; DUF227.
DR Pfam; PF02958; DUF227; 1.
DR SMART; SM00587; CHK; 1.
DR SQU SEQUENCE 394 AA; 44297 MW; D4424A97AE24451 CRC64;

Query Match 38.3%; Score 54; DB 5; Length 394;
Best Local Similarity 60.0%; Pred. No. 6.5;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 11 HPPLPKDIVRSIAKC 25
Db |||:::|||||
62 HPNLPKXNVVWIKIASC 76

RESULT 3
Q94FR6 PRELIMINARY; PRT; 1078 AA.
ID Q94FR6
AC Q94FR6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Acetyl-CoA carboxylase (Fragment).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RA Zagnitko O.; Jelenka J.; Tevzadze G.; Haselkorn R.; Gornicki P.;
RT "An isoleucine/leucine residue in the carboxyltransferase domain of
RT acetyl-CoA carboxylase is critical for interaction with
RT arylxyphenoxypionate and cyclohexanedione inhibitors.";
Proc. Natl. Acad. Sci. U.S.A. 98:6617-6622 (2001).
DR EMBL; AF359518; AAK62314.1; -.
DR GO; GO:0009343; C-biotin carboxylase complex; IEA.
DR GO; GO:0004075; F-biotin carboxylase activity; IEA.
DR InterPro; IPR000022; Carboxyl_trans.
DR Pfam; PF01039; Carboxyl_trans; 1.
DR NON TER 1
FT NON TER 1078
SQ SEQUENCE 1078 AA; 120645 MW; 0ACFC8BE8D794CD1 CRC64;

Query Match 38.3%; Score 54; DB 10; Length 1078;
Best Local Similarity 55.6%; Pred. No. 19;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 11 HPPLPKDIVRSIAKC 25
Db |||:::|||||
62 HPNLPKXNVVWIKIASC 76

RESULT 3
Q94FR6 PRELIMINARY; PRT; 1078 AA.
ID Q94FR6
AC Q94FR6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Acetyl-CoA carboxylase (Fragment).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RA Zagnitko O.; Jelenka J.; Tevzadze G.; Haselkorn R.; Gornicki P.;
RT "An isoleucine/leucine residue in the carboxyltransferase domain of
RT acetyl-CoA carboxylase is critical for interaction with
RT arylxyphenoxypionate and cyclohexanedione inhibitors.";
Proc. Natl. Acad. Sci. U.S.A. 98:6617-6622 (2001).
DR EMBL; AF359518; AAK62314.1; -.
DR GO; GO:0009343; C-biotin carboxylase complex; IEA.
DR GO; GO:0004075; F-biotin carboxylase activity; IEA.
DR InterPro; IPR000022; Carboxyl_trans.
DR Pfam; PF01039; Carboxyl_trans; 1.
DR NON TER 1
FT NON TER 1078
SQ SEQUENCE 1078 AA; 120645 MW; 0ACFC8BE8D794CD1 CRC64;

Query Match 38.3%; Score 54; DB 10; Length 1078;
Best Local Similarity 55.6%; Pred. No. 19;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 3 ETVYSRVTHPLPKDIVR 20
Db |||:::|||||
350 ETVYSRLYQPHLVKDSIQ 367

RESULT 4
Q94FR7 PRELIMINARY; PRT; 1079 AA.
ID Q94FR7
AC Q94FR7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Acetyl-CoA carboxylase (Fragment).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CV. DK592;
RX MEDLINE=21287275; PubMed=11381131;
RT Zagnitko O.; Jelenka J.; Tevzadze G.; Haselkorn R.; Gornicki P.;
RT "An isoleucine/leucine residue in the carboxyltransferase domain of
RT acetyl-CoA carboxylase is critical for interaction with
RT arylxyphenoxypionate and cyclohexanedione inhibitors.";
Proc. Natl. Acad. Sci. U.S.A. 98:6617-6622 (2001).
DR EMBL; AF359517; AAK62313.1; -.
DR GO; GO:0009343; C-biotin carboxylase complex; IEA.
DR GO; GO:0004075; F-biotin carboxylase activity; IEA.
DR InterPro; IPR000022; Carboxyl_trans.
DR Pfam; PF01039; Carboxyl_trans; 1.
DR NON TER 1
FT NON TER 1079
SQ SEQUENCE 1079 AA; 120717 MW; DAD7508EF9BE3AE1 CRC64;

Query Match 38.3%; Score 54; DB 10; Length 1079;
Best Local Similarity 55.6%; Pred. No. 19;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 ETVYSRVTHPLPKDIVR 20
Db |||:::|||||
350 ETVYSRLYQPHLVKDSIQ 367

RESULT 5
Q08367 PRELIMINARY; PRT; 1625 AA.
ID Q08367
AC Q08367;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Acetyl-CoA carboxylase (EC 6.4.1.2) (ACC) (Contains: biotin
DE carboxylase [EC 6.3.4.14]) (Fragment).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=B73; TISSUE=LEAF;
RX MEDLINE=94154242; PubMed=7906562;
RA Ashton A.R.; Jenkins C.L.; Whitfield P.R.;
RT "Molecular cloning of two different cDNAs for maize acetyl CoA
RT carboxylase.";
Plant Mol. Biol. 24:35-49 (1994).
CC -1- FUNCTION: THIS PROTEIN CARRIES THREE CATALYTIC FUNCTIONS: BIOTIN
CC CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE, AND
CC CARBOXYLTRANSFERASE.
CC -1- CATALYTIC ACTIVITY: ATP + ACETYL-COA + HCO(3) (-) = ADP +
CC ORTHOPHOSPHATE + MALONYL-COA.

```

CC -1- COFACTOR: BIOTIN.  
CC -1- ENZYME REGULATION: BY PHOSPHORYLATION.  
CC -1- PATHWAY: FIRST STEP (RATE LIMITING) IN LONG-CHAIN FATTY ACID  
CC SYNTHESIS.  
CC -1- SUBUNIT: HOMOTETRAMER.  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -1- SIMILARITY: PARTIAL TO CARBAMYL PHOSPHATE SYNTHETASES.  
DR EMBL; Z24449; CAA80822.1; -  
DR PIR; S42659; S42659.  
DR GO; GO:0009343; C:biotin carboxylase complex; IEA.  
DR GO; GO:0003989; F:acetyl-CoA carboxylase activity; IEA.  
DR GO; GO:000374; F:biotin binding; IEA.  
DR GO; GO:0004075; F:biotin carboxylase activity; IEA.  
DR GO; GO:0016874; F:ligase activity; IEA.  
DR GO; GO:0006633; F:fatty acid biosynthesis; IEA.  
DR InterPro; IPR000089; Biotin lipoyl.  
DR InterPro; IPR000022; Carboxyl trans.  
DR Pfam; PF00364; biotin lipoyl; 1.  
DR Pfam; PF01039; Carboxyl trans; 1.  
KW Fatty acid biosynthesis; Biotin; Ligase; Multifunctional enzyme.  
FT NON TER 1  
SQ SEQUENCE 1625 AA; 181392 MW; 3624EB042D1DF6B0 CRC64;

Query Match 38.3%; Score 54; DB 10; Length 1625;  
Best Local Similarity 55.6%; Pred. No. 30;  
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 ETYSRVTHPHLPKDIVR 20  
DB 459 ETYISRLYQPHLVKDSIQ 476

## RESULT 6

Q48959 ID O48959 PRELIMINARY; PRT; 2311 AA.  
AC O48959;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Acetyl-coenzyme A carboxylase.  
GN ACC-1.  
OS Triticum aestivum (Wheat).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
OC Triticeae; Triticum.  
OX NCBI\_TaxID=4565;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. HARD RED WINTER TAM 107;  
RX MEDLINE=98054381; PubMed=9391173;  
RA Gornicki P., Paris J., Podkowinski J., Gill B., Haselkorn R.;  
RT "Plastid-localized acetyl-CoA carboxylase of bread wheat is encoded by  
a single gene on each of the three ancestral chromosome sets";  
Proc. Natl. Acad. Sci. U.S.A. 94:14179-14184 (1997).  
DR EMBL; AF029895; AAC39330.1; -  
DR EMBL; AF029896; AAC39331.1; -  
DR PIR; T06161; T06161.  
DR HSSP; P24182; 1DV1.  
DR GO; GO:0009343; C:biotin carboxylase complex; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004075; F:biotin carboxylase activity; IEA.  
DR GO; GO:0016874; F:ligase activity; IEA.  
DR GO; GO:0008152; P:metabolism; IEA.  
DR InterPro; IPR005482; Biotin carb. C.  
DR InterPro; IPR000089; Biotin lipoyl.  
DR InterPro; IPR000022; Carboxyl trans.  
DR InterPro; IPR005479; C:ase L D2.  
DR InterPro; IPR005481; C:ase L N.  
DR Pfam; PF02785; biotin carb. C; 1.  
DR Pfam; PF00364; biotin lipoyl; 1.  
DR Pfam; PF01039; carboxyl trans; 1.  
DR Pfam; PF00289; CFSase\_L chain; 1.

DR Pfam; PF02786; CFSase\_L D2; 1.  
DR PROSITE; PS00867; CFSASE\_2; 1.  
SQ SEQUENCE 2311 AA; 254970 MW; 1BEC5CDD0F9F984B CRC64;

Query Match 38.3%; Score 54; DB 10; Length 2311;  
Best Local Similarity 55.6%; Pred. No. 43;  
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 ETYSRVTHPHLPKDIVR 20  
DB 1159 ETYISRLYQPHLVKDSIQ 1176

## RESULT 7

Q7XYR4 ID Q7XYR4 PRELIMINARY; PRT; 2324 AA.  
AC Q7XYR4;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Acetyl-coenzyme A carboxylase ACC1A.  
GN ACC1A.  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Zong31; TISSUE=Callus;  
RA Zhao H., Wang G.;  
RT "Cloning of chloroplast acetyl-CoA carboxylase cDNA from Zea mays  
callus";  
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY312171; AAP78896.1; -  
DR EMBL; AY312171; AAP78896.1; -  
SQ SEQUENCE 2324 AA; 257167 MW; 4E2249B04E7066A6 CRC64;

Query Match 38.3%; Score 54; DB 10; Length 2324;  
Best Local Similarity 55.6%; Pred. No. 43;  
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 ETYSRVTHPHLPKDIVR 20  
DB 1158 ETYISRLYQPHLVKDSIQ 1175

## RESULT 8

Q41743 ID Q41743 PRELIMINARY; PRT; 2325 AA.  
AC Q41743;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Acetyl-coenzyme A carboxylase.  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A188; TISSUE=Leaf;  
RX MEDLINE=95357420; PubMed=7630949;  
RA Egli M.A., Lutz S.M., Somers D.A., Gengenbach B.G.;  
RT "A maize acetyl-coenzyme A carboxylase cDNA sequence";  
Plant Physiol. 108:1259-1300 (1995).  
DR EMBL; U19183; AAA80214.1; -  
DR PIR; S42659; S42659.  
DR PIR; T02235; T02235.  
DR HSSP; P24182; 1DV1.  
DR GO; GO:0003943; C:biotin carboxylase complex; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004075; F:biotin carboxylase activity; IEA.





DT 01-MAR-2003 (TRENBLrel. 23, Created)  
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
 DE Alpha-fetoprotein precursor.  
 GN AFP.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Body;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs."  
 RL Nature 420:563-573(2002).  
 DR MBL; AK076053; BAC36150.1; -.  
 DR MGD; MGI:87951; Afp.  
 DR GO; GO:0005615; C:extracellular space; IDA.  
 DR GO; GO:0001542; P:ovulation (sensu Mammalia); IMP.  
 DR GO; GO:0042448; P:progesterone metabolism; IMP.  
 DR GO; GO:0019953; P:sexual reproduction; IMP.  
 DR InterPro; IPR000264; Serum\_albumin.  
 DR Pfam; PF00273; transport prot.; 3.  
 DR PRINTS; PR00802; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum\_albumin; 1.  
 DR SMART; SM00103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 2.  
 SQ SEQUENCE 605 AA; 67322 MW; 048B7A4A8B01E4B CRC64;  
  
 Query Match 36.5%; Score 51.5; DB 11; Length 605;  
 Best Local Similarity 57.9%; Pred. No. 25;  
 Matches 11; Conservative 3; Mismatches 4; Indels 1; Gaps 1;  
  
 QY 6 YSRVTHPLPKDIVRSIAK 24  
 ||| |||: ||| : |||  
 Db 354 YSR-THPNLPVSVILRIAK 371  
  
 RESULT 13  
 Q8BK56 PRELIMINARY; PRT; 605 AA.  
 ID Q8BK56  
 AC Q8BK56  
 DT 01-MAR-2003 (TRENBLrel. 23, Created)  
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
 DE Alpha-fetoprotein precursor.  
 GN AFP.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Placenta, and Extraembryonic tissue;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs."  
 RL Nature 420:563-573(2002).  
 DR MBL; AK076197; BAC36249.1; -.  
 DR MGD; MGI:87951; Afp.  
 DR GO; GO:0005615; C:extracellular space; IDA.  
 DR GO; GO:0001542; P:ovulation (sensu Mammalia); IMP.  
 DR GO; GO:0042448; P:progesterone metabolism; IMP.  
 DR GO; GO:0019953; P:sexual reproduction; IMP.  
 DR InterPro; IPR000264; Serum\_albumin.  
 DR Pfam; PF00273; transport prot.; 3.  
 DR PRINTS; PR00802; SERUMALBUMIN.  
 DR ProDom; PD002486; Serum\_albumin; 1.

DR SMART; SM00103; ALBUMIN; 3.  
 DR PROSITE; PS00212; ALBUMIN; 2.  
 SQ SEQUENCE 605 AA; 67409 MW; CF3509A9EC146110 CRC64;  
  
 Query Match 36.5%; Score 51.5; DB 11; Length 605;  
 Best Local Similarity 57.9%; Pred. No. 25;  
 Matches 11; Conservative 3; Mismatches 4; Indels 1; Gaps 1;  
  
 QY 6 YSRVTHPLPKDIVRSIAK 24  
 ||| |||: ||| : |||  
 Db 354 YSR-THPNLPVSVILRIAK 371  
  
 RESULT 14  
 Q7XR27 PRELIMINARY; PRT; 723 AA.  
 ID Q7XR27  
 AC Q7XR27  
 DT 01-OCT-2003 (TRENBLrel. 25, Created)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
 DE OSJNB0014F04.20 protein.  
 GN OSJNB0014F04.20.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Fu G., Wang S.Y., Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F.,  
 RA Jia J., Yin H.F., Zhang Y., Cai Z., Chen X., Kang H., Chen X.Y.,  
 RA Shao Y., Sun Y., Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W.,  
 RA Sheng H.H., Gu J.L., Chen S.T., Ni L., Zhu F.H., Han B., Feng Q.,  
 RA Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X., Liu Y.L., Mu J., Yu Z.,  
 RA Chen L., Fan D.L., Weng Q.J., Zhang L., Lu Y.Q., Yu S.L., Liu X.H.,  
 RA Lu T.T., Zhang Y.J., Lu Y., Li C., Li T., Zhang Y., Hu H., Jia P.X.,  
 RA Qian Y.M., Ying K., Zhou B., Chen Z.H., Hao P., Zhang L., Wu M.,  
 RA Zhang R.Q., Guan J.P., Hong G.F.;  
 RA Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL060449; CAE02854.1; -.  
 SQ SEQUENCE 723 AA; 80810 MW; 2907DBAB6A3B7DA1 CRC64;  
  
 Query Match 36.5%; Score 51.5; DB 10; Length 723;  
 Best Local Similarity 44.4%; Pred. No. 30;  
 Matches 16; Conservative 1; Mismatches 6; Indels 13; Gaps 2;  
  
 QY 3 ETVY-SRVTHPLPK-----DIVRSIAK 25  
 ||| |||: ||| : |||  
 Db 567 QTVYVSRVPHLVANKGSEFWWDITSLSDIFRGIAC 602  
  
 RESULT 15  
 Q9VIA1 PRELIMINARY; PRT; 195 AA.  
 ID Q9VIA1  
 AC Q9VIA1  
 DT 01-MAY-2000 (TRENBLrel. 13, Created)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
 DE Hypothetical protein PYRAB05260.  
 GN PYRAB05260 OR PAB0359.  
 OS Pyrococcus abyssi.  
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
 OC Pyrococcus.  
 OX NCBI\_TaxID=29292;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GES / Orsay;  
 RA Heilig R.;  
 RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome  
 RT structure and evolution."  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ248284; CAB49448.1; -.  
 DR PIR; A75171; A75171.



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 08:27:32 ; Search time 26.2 Seconds  
(without alignments)  
201.482 Million cell updates/sec

Title: US-09-701-623C-6

Perfect score: 141

Sequence: 1 GETTYSRVTHPLPKDIVRSIAK 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	118	83.7	114	14	US-10-152-190-1
2	118	83.7	115	14	US-10-152-190-3
3	118	83.7	117	14	US-10-152-190-2
4	118	83.7	341	9	US-09-401-636-11
5	118	83.7	341	14	US-10-176-664-11
6	118	83.7	346	14	US-10-152-190-10
7	118	83.7	347	14	US-10-152-190-12
8	118	83.7	348	14	US-10-152-190-11
9	118	83.7	426	14	US-10-214-524-28
10	108	76.6	44	14	US-10-076-674-10
11	108	76.6	44	15	US-10-355-161A-10
12	108	76.6	45	14	US-10-076-674-11
13	108	76.6	45	15	US-10-355-161A-11
14	98	69.5	431	10	US-09-479-614-14
15	98	69.5	431	15	US-10-409-772-14

16	98	69.5	496	10	US-09-479-614-2	Sequence 2, Appli
17	98	69.5	496	10	US-09-479-614-29	Sequence 29, Appli
18	98	69.5	496	14	US-10-214-524-25	Sequence 25, Appli
19	98	69.5	496	15	US-10-409-772-2	Sequence 2, Appli
20	98	69.5	496	15	US-10-409-772-29	Sequence 29, Appli
21	91	64.5	345	9	US-09-401-636-10	Sequence 10, Appli
22	91	64.5	345	14	US-10-176-664-10	Sequence 10, Appli
23	91	64.5	567	14	US-10-214-524-33	Sequence 4, Appli
24	89	63.1	25	9	US-09-938-700-4	Sequence 4, Appli
25	88	62.4	426	14	US-10-214-524-26	Sequence 4, Appli
26	85	60.3	115	14	US-10-152-190-4	Sequence 6, Appli
27	85	60.3	129	14	US-10-152-190-6	Sequence 2, Appli
28	85	60.3	222	9	US-09-809-746-2	Sequence 6, Appli
29	85	60.3	222	10	US-09-809-715-6	Sequence 2, Appli
30	85	60.3	320	10	US-09-847-208-6	Sequence 2, Appli
31	85	60.3	323	9	US-09-949-375A-2	Sequence 6, Appli
32	85	60.3	323	9	US-09-949-375A-4	Sequence 6, Appli
33	85	60.3	323	9	US-09-949-375A-6	Sequence 10, Appli
34	85	60.3	330	9	US-09-949-375A-10	Sequence 1, Appli
35	85	60.3	331	9	US-09-401-636-1	Sequence 1, Appli
36	85	60.3	331	14	US-10-176-664-1	Sequence 329, App
37	85	60.3	331	14	US-10-207-855-329	Sequence 8, Appli
38	85	60.3	336	9	US-09-949-375A-8	Sequence 8, Appli
39	85	60.3	342	14	US-10-176-664-8	Sequence 13, Appli
40	85	60.3	342	14	US-10-176-664-8	Sequence 13, Appli
41	85	60.3	347	14	US-10-152-190-13	Sequence 5, Appli
42	85	60.3	427	10	US-09-847-208-5	Sequence 1, Appli
43	85	60.3	428	9	US-09-916-230-1	Sequence 1, Appli
44	85	60.3	428	9	US-09-949-375A-1	Sequence 1, Appli
45	85	60.3	428	13	US-10-047-542-60	Sequence 60, Appli

ALIGNMENTS

RESULT 1

US-10-152-190-1  
; Sequence 1, Application US/10152190  
; Publication No. US20030096369A1  
; GENERAL INFORMATION:  
; APPLICANT: Morsey, Mohamed A.  
; TITLE OF INVENTION: No. US20030096369A1-anaphylactogenic IGE vaccines  
; FILE REFERENCE: PC11011A  
; CURRENT APPLICATION NUMBER: US/10/152,190  
; CURRENT FILING DATE: 2002-05-21  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 114  
; TYPE: PRT  
; ORGANISM: Dog CH3 domain  
US-10-152-190-1

Query Match 83.7%; Score 118; DB 14; Length 114;  
Best Local Similarity 95.7%; Pred No. 2, 2e-09;  
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GETTYSRVTHPLPKDIVRSIAK 24  
DB 84 GETTYSRVTHPLPKDIVRSIAK 106  
|||||

RESULT 2

US-10-152-190-3  
; Sequence 3, Application US/10152190  
; Publication No. US20030096369A1  
; GENERAL INFORMATION:  
; APPLICANT: Morsey, Mohamed A.  
; TITLE OF INVENTION: No. US20030096369A1-anaphylactogenic IGE vaccines  
; FILE REFERENCE: PC11011A  
; CURRENT APPLICATION NUMBER: US/10/152,190  
; CURRENT FILING DATE: 2002-05-21  
; NUMBER OF SEQ ID NOS: 28

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Human CH3/dog CH3 domain chimera
US-10-152-190-3

Query Match      83.7%; Score 118; DB 14; Length 115;
Best Local Similarity 95.7%; Pred. No. 2.2e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 GETTYSRVTHPLPKDIVRSIAK 24
      ||||| ||||| ||||| ||||| |||||
Db      85 GETTICRVTHPLPKDIVRSIAK 107

RESULT 3
US-10-152-190-2
; Sequence 2, Application US/10152190
; Publication No. US20030096369A1
; GENERAL INFORMATION:
; APPLICANT: Morsey, Mohamed A.
; TITLE OF INVENTION: ENHANCED VACCINES
; FILE REFERENCE: PC11011A
; CURRENT APPLICATION NUMBER: US/10/152,190
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Human CH3/dog CH3 domain fusion
US-10-152-190-2

Query Match      83.7%; Score 118; DB 14; Length 117;
Best Local Similarity 95.7%; Pred. No. 2.2e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 GETTYSRVTHPLPKDIVRSIAK 24
      ||||| ||||| ||||| ||||| |||||
Db      87 GETTICRVTHPLPKDIVRSIAK 109

RESULT 4
US-09-401-636-11
; Sequence 11, Application US/09401636
; Patent No. US20010038943A1
; GENERAL INFORMATION:
; APPLICANT: Hellman, Lars T.
; TITLE OF INVENTION: ENHANCED VACCINES
; FILE REFERENCE: 10223/006001
; CURRENT APPLICATION NUMBER: US/09/401,636
; CURRENT FILING DATE: 1999-09-22
; PRIOR APPLICATION NUMBER: US 60/106,652
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated proteins
US-09-401-636-11

Query Match      83.7%; Score 118; DB 9; Length 341;
Best Local Similarity 95.7%; Pred. No. 7.2e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 GETTYSRVTHPLPKDIVRSIAK 24
      ||||| ||||| ||||| ||||| |||||
Db      206 GETTICRVTHPLPKDIVRSIAK 228

RESULT 5
US-10-176-664-11
; Sequence 11, Application US/10176664
; Publication No. US20030031663A1
; GENERAL INFORMATION:
; APPLICANT: Hellman, Lars T.
; TITLE OF INVENTION: ENHANCED VACCINES
; FILE REFERENCE: 10223/006001
; CURRENT APPLICATION NUMBER: US/10/176,664
; CURRENT FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US/09/401,636
; PRIOR FILING DATE: 1999-09-22
; PRIOR APPLICATION NUMBER: US 60/106,652
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated proteins
US-10-176-664-11

Query Match      83.7%; Score 118; DB 14; Length 341;
Best Local Similarity 95.7%; Pred. No. 7.2e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 GETTYSRVTHPLPKDIVRSIAK 24
      ||||| ||||| ||||| ||||| |||||
Db      206 GETTICRVTHPLPKDIVRSIAK 228

RESULT 6
US-10-152-190-10
; Sequence 10, Application US/10152190
; Publication No. US20030096369A1
; GENERAL INFORMATION:
; APPLICANT: Morsey, Mohamed A.
; TITLE OF INVENTION: NO. US20030096369A1-anaphylactogenic IgE vaccines
; FILE REFERENCE: PC11011A
; CURRENT APPLICATION NUMBER: US/10/152,190
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 346
; TYPE: PRT
; ORGANISM: IgE-1 fusion protein
US-10-152-190-10

Query Match      83.7%; Score 118; DB 14; Length 346;
Best Local Similarity 95.7%; Pred. No. 7.3e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 GETTYSRVTHPLPKDIVRSIAK 24
      ||||| ||||| ||||| ||||| |||||
Db      212 GETTICRVTHPLPKDIVRSIAK 234

RESULT 7
US-10-152-190-12
; Sequence 12, Application US/10152190
; Publication No. US20030096369A1
; GENERAL INFORMATION:
; APPLICANT: Morsey, Mohamed A.
; TITLE OF INVENTION: NO. US20030096369A1-anaphylactogenic IgE vaccines
; FILE REFERENCE: PC11011A
; CURRENT APPLICATION NUMBER: US/10/152,190
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
```

Db 289 GETTYCRVTHPHLPKDIVRSIAK 311

APPLICANT: Sokoll, Kenneth K.  
TITLE OF INVENTION: Stabilized Synthetic Immunogen Delivery System  
FILE REFERENCE: Immunogen Delivery System

```
; CURRENT APPLICATION NUMBER: US/10/076,674
; CURRENT FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (20)..(20)
; OTHER INFORMATION: Xaa indicates epsilon-Lys
;
US-10-076-674-11

Query Match          76.6%; Score 108; DB 14; Length 45;
Best Local Similarity 72.0%; Pred. No. 2.2e-08;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYSRVTHPHLPKDIVRSIAKC 25
Db 21 CGETYQSRVTHPHLPALMRSTTKC 45

RESULT 13
US-10-355-161A-11
; Sequence 11, Application US/10355161A
; Publication No. US2004009897A1
; GENERAL INFORMATION:
; APPLICANT: Sokoll, Kenneth K.
; TITLE OF INVENTION: Stabilized Synthetic Immunogen Delivery System
; FILE REFERENCE: Immunogen Delivery System
; CURRENT APPLICATION NUMBER: US/10/355,161A
; CURRENT FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: US 10/076674
; PRIOR FILING DATE: 2002-02-14
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (20)..(20)
; OTHER INFORMATION: Xaa indicates epsilon-Lys
;
US-10-355-161A-11

Query Match          76.6%; Score 108; DB 15; Length 45;
Best Local Similarity 72.0%; Pred. No. 2.2e-08;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYSRVTHPHLPKDIVRSIAKC 25
Db 21 CGETYQSRVTHPHLPALMRSTTKC 45

RESULT 14
US-09-479-614-14
; Sequence 14, Application US/09479614
; Publication No. US20030013183A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine
; APPLICANT: Weber, Eric
; TITLE OF INVENTION: Feline Immunoglobulin E Molecules and Related Methods
; FILE REFERENCE: P-1047
; CURRENT APPLICATION NUMBER: US/09/479,614
; CURRENT FILING DATE: 2000-01-07
; EARLIER APPLICATION NUMBER: 60/115,033
; EARLIER FILING DATE: 1999-01-07
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 14
; LENGTH: 431
```

```
; TYPE: PRT
; ORGANISM: Felis catus
US-09-479-614-14

Query Match          69.5%; Score 98; DB 10; Length 431;
Best Local Similarity 82.6%; Pred. No. 7e-06;
Matches 19; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 GETYYSRVTHPHLPKDIVRSIAK 24
Db 294 GETYQCKVTHPDLPKDIVRSIAK 316

RESULT 15
US-10-409-772-14
; Sequence 14, Application US/10409772
; Publication No. US20030216565A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine
; APPLICANT: Weber, Eric
; TITLE OF INVENTION: Feline Immunoglobulin E Molecules and Related Methods
; FILE REFERENCE: P-1047
; CURRENT APPLICATION NUMBER: US/10/409,772
; CURRENT FILING DATE: 2003-04-07
; PRIOR APPLICATION NUMBER: US/09/479,614
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 14
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Felis catus
US-10-409-772-14

Query Match          69.5%; Score 98; DB 15; Length 431;
Best Local Similarity 82.6%; Pred. No. 7e-06;
Matches 19; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 GETYYSRVTHPHLPKDIVRSIAK 24
Db 294 GETYQCKVTHPDLPKDIVRSIAK 316

Search completed: February 26, 2004, 08:46:54
Job time : 27.2 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 08:19:31 ; Search time 12.6 Seconds  
(without alignments)  
102.432 Million cell updates/sec

Title: US-09-701-623C-6

Perfect score: 141

Sequence: 1 CGETYSRVTHPLPKDIVRSIAK 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:\*
- 1: /cgn2\_6/ptodata/2/iaa/5A COMB.pcp.\*
  - 2: /cgn2\_6/ptodata/2/iaa/5B COMB.pcp.\*
  - 3: /cgn2\_6/ptodata/2/iaa/6A COMB.pcp.\*
  - 4: /cgn2\_6/ptodata/2/iaa/6B COMB.pcp.\*
  - 5: /cgn2\_6/ptodata/2/iaa/PCTUS COMB.pcp.\*
  - 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	118	83.7	426	1	US-08-336-583-2
2	118	83.7	426	5	PCT-US95-13795-2
3	108	76.6	25	3	US-09-100-414B-95
4	108	76.6	25	3	US-09-303-323-95
5	108	76.6	25	4	US-09-770-014-95
6	108	76.6	42	3	US-09-100-414B-98
7	108	76.6	42	3	US-09-100-414B-99
8	108	76.6	42	3	US-09-100-414B-100
9	108	76.6	42	3	US-09-303-323-98
10	108	76.6	42	3	US-09-303-323-99
11	108	76.6	42	3	US-09-303-323-100
12	108	76.6	42	4	US-09-770-014-98
13	108	76.6	42	4	US-09-770-014-99
14	108	76.6	42	4	US-09-770-014-100
15	108	76.6	45	3	US-09-100-414B-101
16	108	76.6	45	3	US-09-303-323-101
17	108	76.6	45	4	US-09-770-014-101
18	108	76.6	45	3	US-09-100-414B-96
19	108	76.6	45	3	US-09-100-414B-97
20	108	76.6	46	3	US-09-303-323-96
21	108	76.6	46	3	US-09-303-323-97
22	108	76.6	46	4	US-09-770-014-96
23	108	76.6	46	4	US-09-770-014-97
24	108	76.6	63	3	US-09-100-414B-102
25	108	76.6	63	3	US-09-303-323-102
26	108	76.6	63	4	US-09-770-014-102
27	98	69.5	431	4	US-09-479-614-14

28	98	69.5	496	4	US-09-479-614-2	Sequence 2, Appli
29	98	69.5	496	4	US-09-479-614-29	Sequence 29, Appli
30	85	60.3	106	2	US-08-232-539D-54	Sequence 54, Appli
31	85	60.3	113	2	US-08-232-539D-56	Sequence 56, Appli
32	74.5	52.8	119	2	US-08-464-025A-1	Sequence 1, Appli
33	72	51.1	22	2	US-08-232-539D-19	Sequence 19, Appli
34	72	51.1	24	2	US-08-232-539D-20	Sequence 20, Appli
35	72	51.1	56	2	US-08-232-539D-18	Sequence 18, Appli
36	67.5	47.9	109	4	US-08-466-163B-1	Sequence 1, Appli
37	67.5	47.9	109	4	US-08-802-096-1	Sequence 1, Appli
38	57	40.4	118	3	US-08-466-151-1	Sequence 1, Appli
39	54	38.3	1313	2	US-08-244-537-2	Sequence 2, Appli
40	54	38.3	2311	4	US-08-934-386-9	Sequence 9, Appli
41	54	38.3	2325	3	US-08-417-089-6	Sequence 6, Appli
42	54	38.3	2325	3	US-08-695-651-6	Sequence 6, Appli
43	54	38.3	2325	3	US-08-930-285-6	Sequence 6, Appli
44	54	38.3	2325	3	US-08-695-421-6	Sequence 6, Appli
45	54	38.3	2325	4	US-08-697-826A-10	Sequence 10, Appli

ALIGNMENTS

RESULT 1  
US-08-336-583-2  
; Sequence 2, Application US/08336583  
; Patent No. 5629415  
; GENERAL INFORMATION:  
; APPLICANT: HOLLIS, GREGORY F.  
; APPLICANT: PATEL, MAYUR D.  
; TITLE OF INVENTION: DNA ENCODING CANINE IMMUNOGLOBULIN E  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CHRISTINE E. CARTY  
; STREET: 126 E. LINCOLN AVENUE  
; CITY: RAHWAY  
; STATE: NEW JERSEY  
; COUNTRY: USA  
; ZIP: 07065-0900

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/336,583  
FILING DATE: 09-NOV-1994  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: CARTY, CHRISTINE E.  
REGISTRATION NUMBER: 36,099  
REFERENCE/DOCKET NUMBER: 19211  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 594-6734  
TELEFAX: (908) 594-4720  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 426 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-336-583-2

Query Match 83.7%; Score 118; DB 1; Length 426;  
Best Local Similarity 95.7%; Pred. No. 5e-11;  
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GETTYSRVTHPLPKDIVRSIAK 24  
DB 289 GETTYSRVTHPLPKDIVRSIAK 311

RESULT 2  
PCT-US95-13795-2  
; Sequence 2, Application PC/TUS9513795  
; GENERAL INFORMATION:  
; APPLICANT: HOLLIS, GREGORY F.  
; APPLICANT: PATEL, MAYUR D.  
; TITLE OF INVENTION: DNA ENCODING CANINE IMMUNOGLOBULINS  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CHRISTINE E. CARTY  
; STREET: 126 E. LINCOLN AVENUE, P.O. BOX 2000  
; CITY: RAHWAY  
; STATE: NEW JERSEY  
; COUNTRY: USA  
; ZIP: 07065-0907  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatenIn Release #1.0, Version #1.25  
; APPLICATION NUMBER: PCT/US95/13795  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CARTY, CHRISTINE E.  
; REGISTRATION NUMBER: 36,099  
; REFERENCE/DOCKET NUMBER: 19211Y  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908) 594-6734  
; TELEFAX: (908) 594-4720  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 426 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US95-13795-2

Query Match 83.7%; Score 118; DB 5; Length 426;  
Best Local Similarity 95.7%; Pred. No. 5e-11;  
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GETYYSRVTHPLPKDIVRSIAK 24  
|||||  
DB 289 GETYYSRVTHPLPKDIVRSIAK 311

RESULT 3  
US-09-100-414B-95  
; Sequence 95, Application US/09100414B  
; Patent No. 6025468  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Chang Yi  
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE  
; NUMBER OF SEQUENCES: 106  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Morgan & Finnegan, L.L.P.  
; STREET: 345 Park Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10154-0054  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC Windows  
; SOFTWARE: Word 97  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/100,414B  
; FILING DATE: 20-JUNE-1998

; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maria H. Lin  
; REGISTRATION NUMBER: 29,323  
; REFERENCE/DOCKET NUMBER: 1151-4157  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-758-4800  
; TELEFAX: 212-751-6849  
; INFORMATION FOR SEQ ID NO: 95:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 25 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-100-414B-95

Query Match 76.6%; Score 108; DB 3; Length 25;  
Best Local Similarity 72.0%; Pred. No. 7.5e-11;  
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAK 25  
|||||  
DB 1 CGETYYSRVTHPLPKDIVRSIAK 25

RESULT 4  
US-09-303-323-95  
; Sequence 95, Application US/09303323  
; Patent No. 6228987  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Chang Yi  
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE  
; NUMBER OF SEQUENCES: 106  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Morgan & Finnegan, L.L.P.  
; STREET: 345 Park Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10154-0054  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC Windows  
; SOFTWARE: Word 97  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/303,323  
; FILING DATE: 30-APR-1999  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/100,414  
; FILING DATE: 20-JUNE-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maria H. Lin  
; REGISTRATION NUMBER: 29,323  
; REFERENCE/DOCKET NUMBER: 1151-4157  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-758-4800  
; TELEFAX: 212-751-6849  
; INFORMATION FOR SEQ ID NO: 95:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 25 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-303-323-95

Query Match 76.6%; Score 108; DB 3; Length 25;  
Best Local Similarity 72.0%; Pred. No. 7.5e-11;  
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAK 25



```
Db      1 CGETYSRVTHPHLPALMRSTTKC 25
||||| : : : : :
RESULT 5
US-09-770-014-95
; Sequence 95, Application US/09770014
; Patent No. 655282
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/770,014
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-770-014-95

Query Match      76.6%; Score 108; DB 4; Length 25;
Best Local Similarity 72.0%; Pred. No. 7.5e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy      1 CGETYSRVTHPHLPKDIVRSIAKC 25
||||| : : : : :
Db      1 CGETYSRVTHPHLPALMRSTTKC 42
||||| : : : : :
RESULT 6
US-09-100-414B-98
; Sequence 98, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 99:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-99

Query Match      76.6%; Score 108; DB 3; Length 42;
Best Local Similarity 72.0%; Pred. No. 1.4e-10;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy      1 CGETYSRVTHPHLPKDIVRSIAKC 25
||||| : : : : :
Db      18 CGETYSRVTHPHLPALMRSTTKC 42
||||| : : : : :
RESULT 7
US-09-100-414B-99
; Sequence 99, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 99:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-99

Query Match      76.6%; Score 108; DB 3; Length 42;
Best Local Similarity 72.0%; Pred. No. 1.4e-10;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
```

```
Best Local Similarity 72.0%; Pred. No. 1.4e-10; Indels 0; Gaps 0;
Matches 18; Conservative 3; Mismatches 4;

QY 1 CGETYSRVTHPLPKDIVRSIAKC 25
    ||||| ||||| ||||| :|||
Db 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 8
US-09-100-414B-100
; Sequence 100, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 100:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-100

Query Match 76.6%; Score 108; DB 3; Length 42;
Best Local Similarity 72.0%; Pred. No. 1.4e-10;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYSRVTHPLPKDIVRSIAKC 25
    ||||| ||||| ||||| :|||
Db 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 9
US-09-303-323-98
; Sequence 98, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 99:
; SEQUENCE CHARACTERISTICS:
```

```
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-98

Query Match 76.6%; Score 108; DB 3; Length 42;
Best Local Similarity 72.0%; Pred. No. 1.4e-10;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYSRVTHPLPKDIVRSIAKC 25
    ||||| ||||| ||||| :|||
Db 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 10
US-09-303-323-99
; Sequence 99, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 99:
; SEQUENCE CHARACTERISTICS:
```

US-09-770-014-98  
; Sequence 98, Application US/09770014  
; Patent No. 655282  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Chang Yi

```
/ FILING DATE: 20-JUNE-1998
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Maria H. Lin
/ REGISTRATION NUMBER: 29,323
/ REFERENCE/DOCKET NUMBER: 1151-4157
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 212-758-4800
/ TELEFAX: 212-751-6849
/ INFORMATION FOR SEQ ID NO: 99:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 42 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ US-09-770-014-99

Query Match 76.6%; Score 108; DB 4; Length 42;
Best Local Similarity 72.0%; Pred. No. 1.4e-10;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CGETYYSRVTHPLPKDIVRSIAKC 25
Db 18 CGETYYSRVTHPLPALMRSTTKC 42

RESULT 14
US-09-770-014-100
/ Sequence 100, Application US/09770014
/ Patent No. 6559282
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Chang Yi
/ TITLE OF INVENTION: NOVEL LHRH PEPTIDE
/ NUMBER OF SEQUENCES: 106
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Morgan & Finnegan, L.L.P.
/ STREET: 345 Park Avenue
/ CITY: New York
/ STATE: NY
/ COUNTRY: USA
/ ZIP: 10154-0054
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC Windows
/ SOFTWARE: Word 97
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/770,014
/ FILING DATE: 20-JUNE-1998
/ CLASSIFICATION: 424
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Maria H. Lin
/ REGISTRATION NUMBER: 29,323
/ REFERENCE/DOCKET NUMBER: 1151-4157
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 212-758-4800
/ TELEFAX: 212-751-6849
/ INFORMATION FOR SEQ ID NO: 100:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 42 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ US-09-770-014-100

Query Match 76.6%; Score 108; DB 4; Length 42;
Best Local Similarity 72.0%; Pred. No. 1.4e-10;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CGETYYSRVTHPLPKDIVRSIAKC 25
Db 18 CGETYYSRVTHPLPALMRSTTKC 42

RESULT 15
US-09-100-414B-101
/ Sequence 101, Application US/09100414B
/ Patent No. 6025468
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Chang Yi
/ TITLE OF INVENTION: NOVEL LHRH PEPTIDE
/ NUMBER OF SEQUENCES: 106
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Morgan & Finnegan, L.L.P.
/ STREET: 345 Park Avenue
/ CITY: New York
/ STATE: NY
/ COUNTRY: USA
/ ZIP: 10154-0054
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC Windows
/ SOFTWARE: Word 97
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/100,414B
/ FILING DATE: 20-JUNE-1998
/ CLASSIFICATION: 424
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Maria H. Lin
/ REGISTRATION NUMBER: 29,323
/ REFERENCE/DOCKET NUMBER: 1151-4157
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 212-758-4800
/ TELEFAX: 212-751-6849
/ INFORMATION FOR SEQ ID NO: 101:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 45 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ US-09-100-414B-101

Query Match 76.6%; Score 108; DB 3; Length 45;
Best Local Similarity 72.0%; Pred. No. 1.5e-10;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CGETYYSRVTHPLPKDIVRSIAKC 25
Db 21 CGETYYSRVTHPLPALMRSTTKC 45

Search completed: February 26, 2004, 08:29:26
Job time : 13.6 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 08:18:56 ; Search time 10 Seconds  
(without alignments)  
240.479 Million cell updates/sec

Title: US-09-701-623C-7  
Perfect score: 145  
Sequence: 1 CGEGYQSRVDHHPFKPIVRSITK 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_78.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	122	84.1	429	1 EHRT	Ig epsilon chain C
2	98	67.6	107	2 I68730	IgE chain C3 regio
3	98	67.6	107	2 I68726	IgE chain C3 regio
4	98	67.6	388	1 EHMS	Ig epsilon chain C
5	98	67.6	548	2 S38864	Ig epsilon chain C
6	85	58.6	423	1 EHMS	Ig epsilon chain C
7	83	57.2	426	2 I36948	Ig epsilon chain C
8	80	55.2	428	1 EHHU	Ig epsilon chain C
9	51	35.2	245	2 I55951	MHC class II E-bet
10	51	35.2	264	2 A60497	H-2 class II histoc
11	51	35.2	264	2 I0989	class II histocomp
12	49	33.8	345	1 HLCHB4	MHC class I histoc
13	49	33.8	385	2 I51309	major histocompati
14	49	33.8	355	2 T28152	MHC class I histoc
15	49	33.8	430	2 A60531	cell cycle protein
16	49	33.8	549	2 S04845	Ig heavy chain pre
17	49	33.8	1208	2 T23222	hypothetical prote
18	48.5	33.4	567	2 T33400	protein kinase C h
19	48.5	33.4	597	2 T33399	protein kinase C h
20	48.5	33.4	704	1 S60117	protein kinase C (
21	48	33.1	93	2 I54421	MHC RT1.B-beta2 -
22	48	33.1	110	2 S43147	Ig epsilon chain -
23	48	33.1	152	2 S14236	Ig gamma-1 chain C
24	48	33.1	237	2 C60497	H-2 class II histoc
25	48	33.1	266	2 A39260	MHC class II histoc
26	48	33.1	266	2 B39260	MHC class II histoc
27	48	33.1	324	1 GIMS	Ig gamma-1 chain C
28	48	33.1	393	1 GIMSM	Ig gamma-1 chain C
29	48	33.1			

30	48	33.1	444	2 PC4436	monoclonal antibod
31	47.5	32.8	1198	2 T49726	hypothetical prote
32	47	32.4	322	2 PS0019	Ig gamma-2a chain
33	47	32.4	1413	2 T26467	hypothetical prote
34	47	32.4	2254	2 D86215	protein T6D22.14 f
35	46.5	32.1	232	2 AG2525	hypothetical prote
36	46.5	32.1	302	2 E86367	hypothetical prote
37	46.5	32.1	510	2 E97307	probable cardiolip
38	46	31.7	86	2 D84566	hypothetical prote
39	46	31.7	116	2 E64392	desulfoferroxidin
40	46	31.7	238	2 A45544	BoLA-DQ beta-1 - b
41	46	31.7	260	2 I45938	MHC cell surface g
42	46	31.7	277	1 MNH32	32K nonstructural
43	46	31.7	314	2 AB3610	multidrug resistan
44	46	31.7	328	2 T47161	Ig gamma 3 chain c
45	46	31.7	328	2 T47158	Ig gamma 1 chain c

## ALIGNMENTS

### RESULT 1

EHRT

Ig epsilon chain C region - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 17-Dec-1982 #sequence revision 17-Dec-1982 #text\_change 16-Jul-1999

C:Accession: A93442; A90937; A02143

R:Hellman, L.; Pettersson, U.; Engstrom, A.; Karlsson, T.; Bennich, H.

Nucleic Acids Res. 10, 6041-6049, 1982

A:Title: Structure and evolution of the heavy chain from rat immunoglobulin E.

A:Reference number: A93442; MUID:83064537; PMID:6252865

A:Accession: A93442

A:Molecule type: mRNA

A:Residues: 1-429 <HEL>

A:Experimental source: strain LOU/c/Wel, immunocytooma IR2

R:Kindsvogel, W.R.; Reddy, E.P.; Moore, J.M.; Faust Jr., C.H.

DNA J, 335-343, 1982

A:Title: A cloned cDNA probe for rat immunoglobulin epsilon heavy chain: construction,

A:Reference number: A90937; MUID:83182019; PMID:6820340

A:Contents: myeloma IR162

A:Accession: A90937

A:Molecule type: mRNA

A:Residues: 'N', 169-307, 'L', 309-342 <KIN>

C:Complex: An immunoglobulin heterotrimer subunit consists of two identical light (ka

hain disulfide bonds. In some cases, such as IGA and IgM, the subunits associate into 1

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: duplication; glycoprotein; heterotrimer; immunoglobulin

F:19-80/Domain: immunoglobulin homology <IM1>

F:118-186/Domain: immunoglobulin homology <IM2>

F:223-291/Domain: immunoglobulin homology <IM3>

F:327-398/Domain: immunoglobulin homology <IM4>

F:46,99,170,240,265,369,419/Binding site: carbonylde (Asn) (covalent) #status predict

Query Match

Best Local Similarity 84.1%; Score 122; DB 1; Length 429;

Mismatches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 CGEGYQSRVDHHPFKPIVRSITK 24

DB 284 CGEGYQSRVDHHPFKPIVRSITK 306

### RESULT 2

I68730

IgE chain C3 region - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 21-Jan-2000

C:Accession: I68730

R:Shinkai, Y.; Nakauchi, H.; Honjo, T.; Okumura, K.

Immunogenetics 27, 288-292, 1988

A:Title: Mouse immunoglobulin allotypes: Multiple differences between the nucleic acid

A:Reference number: I54443; MUID:88152907; PMID:3346043

A:Accession: I68730

DB 247 GYGQCVVDHPDPFKPIVRSITK 269

RESULT 5

S38864

IG epsilon chain C region - mouse (fragment)

C/Species: Mus musculus (house mouse)

C/Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 24-May-2001

C/Accession: S38864

R/Klipp, B.; Becker, W.; Schlaak, M., November 1993

Submitted to the EMBL Data Library, November 1993

A/Description: Combination of a defined specificity and desired isotype by cloning of a

A/Reference number: S38864

A/Accession: S38864

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-548 <IP>

A/Cross-references: EMBL:Z73797; NID:G416537; PIDN:CAA81788.1; PID:G940782

C/Superfamily: immunoglobulin C region; immunoglobulin homology

F/353-421/Domain: immunoglobulin homology <IMM>

Query Match 67.6%; Score 98; DB 2; Length 548;

Best Local Similarity 82.6%; Pred. No. 1.8e-06;

Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPIVRSITK 24

DB 414 GYGQCVVDHPDPFKPIVRSITK 436

RESULT 6

EMMS

IG epsilon chain C region (version 2) - mouse

C/Species: Mus musculus (house mouse)

C/Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 16-Aug-1996

C/Accession: A02145

R/Ishida, N.; Ueda, S.; Hayashida, H.; Miyata, T.; Honjo, T.

EMBO J. 1, 1117-1123, 1982

A/Title: The nucleotide sequence of the mouse immunoglobulin epsilon gene: comparison w

A/Reference number: A90966; MUID:84236092; PMID:6329728

A/Accession: A02145

A/Molecule type: DNA

A/Residues: 1-423 <ISH>

A/Note: the sequence was determined from the germline gene

C/Genetics:

A/Introns: 91/1; 199/1; 307/1

C/Complex: An immunoglobulin heterotetramer subunit consists of two identical light (ka

hain disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into 1

C/Superfamily: immunoglobulin C region; immunoglobulin homology

C/Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin

F/15-77/Domain: immunoglobulin homology <IMM1>

F/115-183/Domain: immunoglobulin homology <IMM2>

F/220-288/Domain: immunoglobulin homology <IMM3>

F/325-396/Domain: immunoglobulin homology <IMM4>

F/23-75,122-181,227-286,332-394/Disulfide bonds: #status predicted

F/43,84,167,239,262,417/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 58.6%; Score 85; DB 1; Length 423;

Best Local Similarity 77.3%; Pred. No. 0.00011;

Matches 17; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPIVRSIT 23

DB 281 GYGQCVVDHPDPFKPIVRSIT 302

RESULT 7

I36948

IG epsilon chain - chimpanzee (fragment)

C/Species: Pan troglodytes (chimpanzee)

C/Date: 04-Oct-1996 #sequence\_revision 04-Oct-1996 #text\_change 21-Jan-2000

C/Accession: I36948

R/Sakoyama, Y.; Hong, K.

A:Molecule type: protein  
A:Residues: 'GAWTL',6,'X',8-16,'B',18-43,'B',45-52,55-92,95-97,'B',99-121,'B',123,'L',125-127,'B',129-130,'B',132-133,'B',135-136,'B',138-139,'B',141-142,'B',144-145,'B',147-148,'B',150-151,'B',153-154,'B',156-157,'B',159-160,'B',162-163,'B',165-166,'B',168-169,'B',171-172,'B',174-175,'B',177-178,'B',180-181,'B',183-184,'B',186-187,'B',189-190,'B',192-193,'B',195-196,'B',198-199,'B',201-202,'B',204-205,'B',207-208,'B',210-211,'B',213-214,'B',216-217,'B',219-220,'B',222-223,'B',225-226,'B',228-229,'B',231-232,'B',234-235,'B',237-238,'B',240-241,'B',243-244,'B',246-247,'B',249-250,'B',252-253,'B',255-256,'B',258-259,'B',261-262,'B',264-265,'B',267-268,'B',270-271,'B',273-274,'B',276-277,'B',279-280,'B',282-283,'B',285-286,'B',288-289,'B',291-292,'B',294-295,'B',297-298,'B',300-301,'B',302-303,'B',305-306,'B',308-309,'B',311-312,'B',314-315,'B',317-318,'B',320-321,'B',323-324,'B',326-327,'B',329-330,'B',332-333,'B',335-336,'B',338-339,'B',341-342,'B',344-345,'B',347-348,'B',350-351,'B',353-354,'B',356-357,'B',359-360,'B',362-363,'B',365-366,'B',368-369,'B',371-372,'B',374-375,'B',377-378,'B',380-381,'B',383-384,'B',386-387,'B',389-390,'B',392-393,'B',395-396,'B',398-399,'B',401-402,'B',404-405,'B',407-408,'B',410-411,'B',413-414,'B',416-417,'B',419-420,'B',422-423,'B',425-426,'B',428-429,'B',431-432,'B',434-435,'B',437-438,'B',440-441,'B',443-444,'B',446-447,'B',449-450,'B',452-453,'B',455-456,'B',458-459,'B',461-462,'B',464-465,'B',467-468,'B',470-471,'B',473-474,'B',476-477,'B',479-480,'B',482-483,'B',485-486,'B',488-489,'B',491-492,'B',494-495,'B',497-498,'B',500-501,'B',503-504,'B',506-507,'B',509-510,'B',512-513,'B',515-516,'B',518-519,'B',521-522,'B',524-525,'B',527-528,'B',530-531,'B',533-534,'B',536-537,'B',539-540,'B',542-543,'B',545-546,'B',548-549,'B',551-552,'B',554-555,'B',557-558,'B',560-561,'B',563-564,'B',566-567,'B',569-570,'B',572-573,'B',575-576,'B',578-579,'B',581-582,'B',584-585,'B',587-588,'B',590-591,'B',593-594,'B',596-597,'B',599-600,'B',602-603,'B',605-606,'B',608-609,'B',611-612,'B',614-615,'B',617-618,'B',620-621,'B',623-624,'B',626-627,'B',629-630,'B',632-633,'B',635-636,'B',638-639,'B',641-642,'B',644-645,'B',647-648,'B',650-651,'B',653-654,'B',656-657,'B',659-660,'B',662-663,'B',665-666,'B',668-669,'B',671-672,'B',674-675,'B',677-678,'B',680-681,'B',683-684,'B',686-687,'B',689-690,'B',692-693,'B',695-696,'B',698-699,'B',701-702,'B',704-705,'B',707-708,'B',710-711,'B',713-714,'B',716-717,'B',719-720,'B',722-723,'B',725-726,'B',728-729,'B',731-732,'B',734-735,'B',737-738,'B',740-741,'B',743-744,'B',746-747,'B',749-750,'B',752-753,'B',755-756,'B',758-759,'B',761-762,'B',764-765,'B',767-768,'B',770-771,'B',773-774,'B',776-777,'B',779-780,'B',782-783,'B',785-786,'B',788-789,'B',791-792,'B',794-795,'B',797-798,'B',800-801,'B',803-804,'B',806-807,'B',809-810,'B',812-813,'B',815-816,'B',818-819,'B',821-822,'B',824-825,'B',827-828,'B',830-831,'B',833-834,'B',836-837,'B',839-840,'B',842-843,'B',845-846,'B',848-849,'B',851-852,'B',854-855,'B',857-858,'B',860-861,'B',863-864,'B',866-867,'B',869-870,'B',872-873,'B',875-876,'B',878-879,'B',881-882,'B',884-885,'B',887-888,'B',890-891,'B',893-894,'B',896-897,'B',899-900,'B',902-903,'B',905-906,'B',908-909,'B',911-912,'B',914-915,'B',917-918,'B',920-921,'B',923-924,'B',926-927,'B',929-930,'B',932-933,'B',935-936,'B',938-939,'B',941-942,'B',944-945,'B',947-948,'B',950-951,'B',953-954,'B',956-957,'B',959-960,'B',962-963,'B',965-966,'B',968-969,'B',971-972,'B',974-975,'B',977-978,'B',980-981,'B',983-984,'B',986-987,'B',989-990,'B',992-993,'B',995-996,'B',998-999,'B',1000-1001,'B',1003-1004,'B',1006-1007,'B',1009-1010,'B',1012-1013,'B',1015-1016,'B',1018-1019,'B',1021-1022,'B',1023-1024,'B',1026-1027,'B',1029-1030,'B',1032-1033,'B',1035-1036,'B',1038-1039,'B',1041-1042,'B',1043-1044,'B',1046-1047,'B',1049-1050,'B',1052-1053,'B',1055-1056,'B',1058-1059,'B',1061-1062,'B',1064-1065,'B',1067-1068,'B',1070-1071,'B',1073-1074,'B',1076-1077,'B',1079-1080,'B',1082-1083,'B',1085-1086,'B',1088-1089,'B',1091-1092,'B',1094-1095,'B',1097-1098,'B',1100-1101,'B',1103-1104,'B',1106-1107,'B',1109-1110,'B',1112-1113,'B',1115-1116,'B',1118-1119,'B',1121-1122,'B',1124-1125,'B',1127-1128,'B',1130-1131,'B',1133-1134,'B',1136-1137,'B',1139-1140,'B',1142-1143,'B',1145-1146,'B',1148-1149,'B',1151-1152,'B',1154-1155,'B',1157-1158,'B',1160-1161,'B',1163-1164,'B',1166-1167,'B',1169-1170,'B',1172-1173,'B',1175-1176,'B',1178-1179,'B',1181-1182,'B',1184-1185,'B',1187-1188,'B',1190-1191,'B',1193-1194,'B',1196-1197,'B',1199-1200,'B',1202-1203,'B',1205-1206,'B',1208-1209,'B',1211-1212,'B',1214-1215,'B',1217-1218,'B',1220-1221,'B',1223-1224,'B',1226-1227,'B',1229-1230,'B',1232-1233,'B',1235-1236,'B',1238-1239,'B',1241-1242,'B',1244-1245,'B',1247-1248,'B',1250-1251,'B',1253-1254,'B',1256-1257,'B',1259-1260,'B',1262-1263,'B',1265-1266,'B',1268-1269,'B',1271-1272,'

Best Local Similarity 60.9%; Pred. No. 0.0006; Indels 0; Gaps 0;  
Matches 14; Conservative 3; Mismatches 6;

QY 2 GEGYQSRVDHPHPKPIVRITK 24

DB 294 GEGYQCRVTHPLPRALMRSTTK 316

## RESULT 9

MHC class II E-beta protein - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 21-Jan-2000  
C:Accession: I55951  
J:Robertson, K.A.; McMaster, W.R.  
J:Immunol. 135, 4095-4099, 1995  
A:Title: Complete structure of a rat RT1 E-beta chain: Extensive conservation of MHC class II beta chain  
A:Reference number: I55951; MUID:86060895; PMID:3934274  
A:Accession: I55951  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-245 <RES>  
A:Cross-references: GB:M1282; NID:G205437; PIDN:AAA41610.1; PID:G205438  
C:Superfamily: class II histocompatibility antigen; immunoglobulin homology  
F:118-183/Domain: immunoglobulin homology <IMM>

Query Match 35.2%; Score 51; DB 2; Length 245;

Best Local Similarity 47.1%; Pred. No. 6.1; Indels 0; Gaps 0;

Matches 8; Conservative 3; Mismatches 6;

QY 2 GEGYQSRVDHPHPKPI 18

DB 176 GEGYTCQVEHPSPV 192

## RESULT 10

A60497  
H-2 class II histocompatibility antigen RT1.D-mu beta chain precursor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 21-Jan-2000  
C:Accession: A60497  
R:Holowachuk, E.W.; Greer, M.K.  
Diabetes 38, 267-271, 1989  
A:Title: Unaltered class II histocompatibility antigens and pathogenesis of IDDM in BB mice  
A:Reference number: A60497; MUID:89121214; PMID:2464510  
A:Accession: A60497  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-264 <HOL>  
C:Superfamily: class II histocompatibility antigen; immunoglobulin homology  
F:137-202/Domain: immunoglobulin homology <IMM>

Query Match 35.2%; Score 51; DB 2; Length 264;

Best Local Similarity 47.1%; Pred. No. 6.6; Indels 0; Gaps 0;

Matches 8; Conservative 3; Mismatches 6;

QY 2 GEGYQSRVDHPHPKPI 18

DB 195 GEGYTCQVEHPSPV 211

## RESULT 11

S10989  
class II histocompatibility antigen RT1-D beta-I chain precursor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 21-Jan-2000  
C:Accession: S10989  
R:Syha-Vedelhauser, J.; Reske, K.  
Nucleic Acids Res. 18, 4598, 1990  
A:Title: Sequence of rat cDNA clone pLR-beta-112 coding for the RT1.D-beta' chain.  
A:Reference number: S10989; MUID:90356406; PMID:2388838  
A:Accession: S10989  
A:Molecule type: mRNA

A:Residues: 1-264 <SVH>  
A:Cross-references: EMBL:X53054; NID:G57169; PIDN:CAA37221.1; PID:G57170  
C:Superfamily: class II histocompatibility antigen; immunoglobulin homology  
C:Keywords: glycoprotein; heterodimer; transmembrane protein  
F:1-26/Domain: signal sequence #status predicted <SIG>  
F:27-264/Product: class II histocompatibility antigen, RT1-D beta-I chain #status predicted  
F:137-202/Domain: immunoglobulin homology <IMM>  
F:129-246/Domain: transmembrane #status predicted <TM>  
F:46/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 35.2%; Score 51; DB 2; Length 264;

Best Local Similarity 47.1%; Pred. No. 6.6; Indels 0; Gaps 0;

Matches 8; Conservative 3; Mismatches 6;

QY 2 GEGYQSRVDHPHPKPI 18

DB 195 GEGYTCQVEHPSPV 211

## RESULT 12

HLCHB4  
MHC class I histocompatibility antigen B-F IV alpha chain precursor - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 22-Jun-1999  
C:Accession: A45846; S01172  
R:Kroemer, G.; Zoorob, R.; Auffray, C.  
Immunogenetics 31, 405-409, 1990  
A:Title: Structure and expression of a chicken MHC class I gene.  
A:Reference number: A45846; MUID:90316612; PMID:2370087  
A:Accession: A45846  
A:Molecule type: DNA  
A:Residues: 1-345 <KRO>  
A:Cross-references: GB:M31012  
R:Guillemot, F.; Billault, A.; Pourquie, O.; Behar, G.; Chausse, A.M.; Zoorob, R.; Kreil, G.  
EMBO J. 7, 2775-2785, 1988  
A:Title: A molecular map of the chicken major histocompatibility complex: the class II region  
A:Reference number: S01172; MUID:89030642; PMID:3141149  
A:Accession: S01172  
A:Molecule type: mRNA  
A:Residues: 1-345 <GUI>  
A:Cross-references: EMBL:X12780; NID:G63089; PIDN:CAA31272.1; PID:G63090  
C:Genetics:  
A:Map position: 16  
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology  
C:Keywords: glycoprotein; heterodimer; transmembrane protein  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-345/Product: class I histocompatibility antigen B-F IV alpha chain #status predicted  
F:23-301/Domain: extracellular #status predicted <EXT>  
F:23-110/Domain: alpha-1 <EXT1>  
F:111-201/Domain: alpha-2 <EXT2>  
F:214-279/Domain: immunoglobulin homology <IMM>  
F:302-324/Domain: transmembrane #status predicted <TM>  
F:325-345/Domain: intracellular #status predicted <INT>  
F:59.107/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:121-183,221-277/Disulfide bonds: #status predicted

Query Match 33.8%; Score 49; DB 1; Length 345;

Best Local Similarity 50.0%; Pred. No. 17; Indels 5; Gaps 0;

Matches 8; Conservative 3; Mismatches 5;

QY 2 GEGYQSRVDHPHPKPI 17

DB 272 GEGYQCRVEHSLPOP 287

## RESULT 13

I51309  
major histocompatibility complex class I glycoprotein haplotype B21 - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 04-Sep-1997 #sequence\_revision 04-Sep-1997 #text\_change 21-Jan-2000  
C:Accession: I51309  
R:Fulton, J.E.; Thacker, E.L.; Bacon, L.D.; Hunt, H.D.  
Eur. J. Immunol. 25, 2069-2076, 1995



A;Title: Functional analysis of avian class I (BFTV) glycoproteins by epitope tagging and  
A;Reference number: I51309; MUID:95347411; PMID:7621880  
A;Accession: I51309  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-355 <FUL>  
A;Cross-references: GB:S78682; NID:g1042200; PIDN:AAB34945.1; PID:g1042201  
C;Genetics:  
A;Gene: BFTV21  
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology  
F;213-278/Domain: immunoglobulin homology <IMM>

Query Match 33.8%; Score 49; DB 2; Length 355;  
Best Local Similarity 50.0%; Pred. No. 18;  
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKP 17  
|:|:|:|:|:|:|:  
Db 271 GDKYQCRVHASLPQP 286

## RESULT 14

T28149  
MHC class I histocompatibility antigen B-F alpha chain 2 - chicken  
C;Species: Gallus gallus (chicken)  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 20-Jun-2000  
C;Accession: T28149  
R;Milne, S.; Kaufman, J.; Beck, S.  
submitted to the EMBL Data Library, May 1998  
A;Description: DNA sequencing and analysis of the chicken major histocompatibility compl  
A;Reference number: Z20475  
A;Accession: T28149  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-355 <MIL>  
A;Cross-references: EMBL:AL023516; PIDN:CAA18969.1  
A;Experimental source: clone CB12  
C;Genetics:  
A;Gene: BFA2  
A;Map position: 16  
A;Introns: 22/1; 110/1; 201/1; 292/1; 328/1; 339/1; 350/1  
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 33.8%; Score 49; DB 2; Length 355;  
Best Local Similarity 50.0%; Pred. No. 18;  
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKP 17  
|:|:|:|:|:|:|:  
Db 271 GDKYQCRVHASLPQP 286

## RESULT 15

T28152  
MHC class I histocompatibility antigen B-F IV alpha chain precursor - chicken  
C;Species: Gallus gallus (chicken)  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 20-Jun-2000  
C;Accession: T28152  
R;Milne, S.; Kaufman, J.; Beck, S.  
submitted to the EMBL Data Library, May 1998  
A;Description: DNA sequencing and analysis of the chicken major histocompatibility compl  
A;Reference number: Z20475  
A;Accession: T28152  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-355 <MIL>  
A;Cross-references: EMBL:AL023516; PIDN:CAA18972.1  
A;Experimental source: clone CB12  
C;Genetics:  
A;Gene: BFA1  
A;Map position: 16  
A;Introns: 22/1; 110/1; 201/1; 292/1; 328/1; 339/1; 350/1  
C;Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 33.8%; Score 49; DB 2; Length 355;  
Best Local Similarity 50.0%; Pred. No. 18;  
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKP 17  
|:|:|:|:|:|:|:  
Db 271 GDKYQCRVHASLPQP 286

Search completed: February 26, 2004, 08:28:14  
Job time : 11 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 08:17:40 ; Search time 6.4 Seconds  
(without alignments)  
203.399 Million cell updates/sec

Title: US-09-701-623C-7

Perfect score: 145

Sequence: 1 CGEGYQSRVDHPHFPKPIVRSITKC 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	122	84.1	429	1 EPC RAT	P01855 rattus norv
2	98	67.6	421	1 EPC MOUSE	P06336 mus musculu
3	80	55.2	428	1 EPC HUMAN	P01854 homo sapien
4	51	35.2	264	1 HB2D RAT	P18211 rattus norv
5	49	33.8	345	1 HALF CHICK	P15979 gallus gall
6	49	33.8	485	1 MURE FUSNN	O8r635 fusbacteri
7	48.5	33.4	704	1 KPCI CAEEL	P34722 caenorhabdi
8	48	33.1	324	1 GCI MOUSE	P01868 mus musculu
9	48	33.1	393	1 GCI MOUSE	P01869 mus musculu
10	47	32.4	322	1 GCA RAT	P20760 rattus norv
11	46.5	32.1	345	1 GHIT HUMAN	O9h3k2 homo sapien
12	46	31.7	116	1 SOR META	O58151 methanococ
13	46	31.7	277	1 VNST CUVO	P18517 bovine coro
14	46	31.7	336	1 GCB MOUSE	P01866 mus musculu
15	46	31.7	370	1 DCUP SCHPO	Q9u5j5 schizosacch
16	46	31.7	373	1 MTF2 MOUSE	Q02395 mus musculu
17	46	31.7	405	1 GCB MOUSE	P01867 mus musculu
18	46	31.7	516	1 MEFA XENLA	Q03414 xenopus lae
19	46	31.7	593	1 MTF2 HUMAN	O9y483 homo sapien
20	45	31.0	116	1 VKI4 YEAST	P36078 saccharomyc
21	45	31.0	330	1 GCAA MOUSE	P01863 mus musculu
22	45	31.0	399	1 GCAN MOUSE	P01865 mus musculu
23	44.5	30.7	387	1 GCP RICPR	Q9ze88 rickettsia
24	44	30.3	104	1 Y4EB RHISN	P55425 rhizobium s
25	44	30.3	252	1 EX12 ARATH	Q9ldj3 arabidopsis
26	44	30.3	263	1 HB2B RAT	P29826 rattus norv
27	44	30.3	275	1 Y760 META	O58170 methanococ
28	44	30.3	290	1 GC3 HUMAN	P01860 homo sapien
29	44	30.3	326	1 GC2 HUMAN	P01859 homo sapien
30	44	30.3	329	1 GC3 MOUSE	P22436 mus musculu
31	44	30.3	330	1 GC1 HUMAN	P01857 homo sapien
32	44	30.3	371	1 HA12 RAT	P16391 rattus norv
33	44	30.3	371	1 RIBD_BACAM	P70814 b riboflavi

34	44	30.3	384	1 GBA2 PEA	Q04279 pisum sativ
35	44	30.3	398	1 GC3M MOUSE	P03987 mus musculu
36	44	30.3	446	1 MUC CHICK	P01875 gallus gall
37	44	30.3	658	1 KPCI LYPTI	Q25378 lytechinus
38	44	30.3	1106	1 ACLY CAEEL	P35585 caenorhabdi
39	43.5	30.0	244	1 TRMB PSSSM	Q88af5 pseudomonas
40	43.5	30.0	303	1 TTI ARATH	Q8vwg3 arabidopsis
41	43.5	30.0	521	1 FXN4 MOUSE	Q8k3g3 mus musculu
42	43.5	30.0	581	1 FZD8 XENLA	O93274 xenopus lae
43	43	29.7	51	1 OC90 CAVPO	P18669 cavia porce
44	43	29.7	119	1 RS20 HUMAN	P17075 homo sapien
45	43	29.7	119	1 RS20 XENLA	P33403 xenopus lae

## ALIGNMENTS

### RESULT 1

ID	EPC_RAT	STANDARD	PRT	429 AA
AC	P01855;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Ig epsilon chain C region.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A. (IMMUNOCYTOMA IR2).			
RC	STRAIN=LOU/C/WSL;			
RX	MEDLINE=83064537; PubMed=6292865;			
RA	Hellman L., Petterson U., Engstrom A., Karlsson T., Bennich H.;			
RT	"Structure and evolution of the heavy chain from rat immunoglobulin E.;"			
RT	Nucleic Acids Res. 10:6041-6049(1982).			
RN	[2]			
RP	SEQUENCE OF 168-342 FROM N.A. (MYELOMA IR162).			
RX	MEDLINE=83182019; PubMed=6820340;			
RA	Kindsvogel W.R., Reddy E.P., Moore J.M., Faust C.H. Jr.;			
RT	"A cloned cDNA probe for rat immunoglobulin epsilon heavy chain: construction, identification, and DNA sequence.;"			
RL	DNA 1:335-343(1982).			
RN	[3]			
RP	SEQUENCE OF 205-306 FROM N.A.			
RX	MEDLINE=82174576; PubMed=6803238;			
RA	Hellman L., Petterson U., Bennich H.;			
RT	"Characterization and molecular cloning of the mRNA for the heavy (epsilon) chain of rat immunoglobulin E.;"			
RL	Proc. Natl. Acad. Sci. U.S.A. 79:1264-1268(1982).			
CC	- - SIMILARITY: Contains 4 immunoglobulin-like domains.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; J00744; AAA41379.1; ALT_INIT.			
DR	PIR; A93442; EHRT.			
DR	HSSP; P01854; 1IGB.			
DR	InterPro; IPR007110; IG-like.			
DR	InterPro; IPR003597; IG_c1.			
DR	InterPro; IPR003006; IG_MHC.			
DR	Pfam; PF00047; ig_4.			
DR	SMART; SM00407; Igcl; 1.			
DR	PROSITE; PS00835; IG_LIKE; 4.			
DR	PROSITE; PS00290; IG_MHC; 3.			
KW	Immunoglobulin domain; Immunoglobulin C region; Repeat.			
FT	NON_TER 1			

```

FT DOMAIN 6 89 IG-LIKE 1.
FT DOMAIN 103 201 IG-LIKE 2.
FT DOMAIN 205 305 IG-LIKE 3.
FT DOMAIN 314 414 IG-LIKE 4.
FT CONFLICT 168 168 R->N (IN REF. 2).
FT CONFLICT 308 308 P->L (IN REF. 2).
SQ SEQUENCE 429 AA; 48671 MW; D2970B34EFA72B0 CRC64;

Query Match 84.1%; Score 122; DB 1; Length 429;
Best Local Similarity 95.7%; Pred. No. 3.2e-10;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GEGYQSRVDHPFKPIVRSITK 24
||||| ||||| ||||| ||||| |||||
Db 284 GEGYQCRVDHPFKPIVRSITK 306

RESULT 2
EPC_MOUSE STANDARD; PRT; 421 AA.
AC P06336; F01856;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig epsilon chain C region.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84236092; PubMed=6329728;
RA Ishida N., Ueda S., Hayashida H., Miyata T., Honjo T.;
RT "The nucleotide sequence of the mouse immunoglobulin epsilon gene:
RT comparison with the human epsilon gene sequence.";
RL EMBO J. 1:1117-1123 (1982).
RN [2]
RP REVISIONS.
RA Honjo T.;
RL Submitted (APR-1986) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 34-421 FROM N.A.
RX MEDLINE=83117774; PubMed=6818553;
RA Liu F.-T., Albrandt K., Sutcliffe J.G., Katz D.H.;
RT "Cloning and nucleotide sequence of mouse immunoglobulin epsilon
RT chain cDNA.";
RN Proc. Natl. Acad. Sci. U.S.A. 79:7852-7856 (1982).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X01857; CAA25977.1; -.
DR EMBL; X01857; CAA25978.1; -.
DR PIR; A02144; EHMS.
DR PIR; A02145; EHMS.
DR HSSP; P01854; LIGE.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003597; IG cl.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00407; IGcl; 2.
DR PROSITE; PS00835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 3.
DR Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON TER 1 90 CH1.
FT DOMAIN 91 197 CH2.
FT DOMAIN 198 304 CH3.

```

```

FT DOMAIN 305 421 CH4.
FT DISULFID 23 75 BY SIMILARITY.
FT DISULFID 121 180 BY SIMILARITY.
FT DISULFID 226 285 BY SIMILARITY.
FT DISULFID 330 392 BY SIMILARITY.
FT CARBOHYD 43 43 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 95 95 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 261 261 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 365 365 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 415 415 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 421 AA; 47320 MW; 8F909EIF30A06B47 CRC64;

Query Match 67.6%; Score 98; DB 1; Length 421;
Best Local Similarity 82.6%; Pred. No. 9.3e-07;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 GEGYQSRVDHPFKPIVRSITK 24
||||| ||||| ||||| ||||| |||||
Db 280 GYGYQCIVDHPFKPIVRSITK 302

RESULT 3
EPC_HUMAN STANDARD; PRT; 428 AA.
AC P01854;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig epsilon chain C region.
GN IGHE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83168997; PubMed=6300763;
RA Seno M., Kurekawa T., Ono Y., Sasada R., Igarashi K.,
RA Kikuchi M., Sugino Y., Nishida Y., Honjo T.;
RT "Molecular cloning and nucleotide sequencing of human immunoglobulin
RT epsilon chain cDNA.";
RL Nucleic Acids Res. 11:719-726 (1983).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT LEU-359.
RX MEDLINE=83001945; PubMed=6288268;
RA Max E.E., Battey J., Ney R., Kirsch I.R., Leder P.;
RT "Duplication and deletion in the human immunoglobulin epsilon genes.";
RL Cell 29:691-699 (1982).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=84236029; PubMed=6234164;
RA Flanagan J.G., Rabbitts T.H.;
RT "The sequence of a human immunoglobulin epsilon heavy chain constant
RT region gene, and evidence for three non-allelic genes.";
RL EMBO J. 1:655-660 (1982).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=84207910; PubMed=6327276;
RA Ueda S., Nakai S., Nishida Y., Hisajima H., Honjo T.;
RT "Long terminal repeat-like elements flank a human immunoglobulin
RT epsilon pseudogene that lacks introns.";
RL EMBO J. 1:1539-1544 (1982).
RN [5]
RP PRELIMINARY SEQUENCE (MYELOMA PROTEIN ND).
RA Bennich H.H., Johansson S.G.O., von Bahr-Lindstrom H.,
RL (In) Bach M.K. (eds.);
RL Immediate hypersensitivity: modern concepts and developments, pp.1-36,
RL Marcel Dekker, New York (1978).
RN [6]

```

RP SEQUENCE OF 1-40; 68-114 AND 427-428 FROM N.A.  
RX MEDLINE=83065234; PubMed=6815656;  
RA Kenten J.H., Molgaard H.V., Houghton M., Derbyshire R.B., Viney J.,  
BA Bell L.O., Gould H.J.;  
BT "Cloning and sequence determination of the gene for the human  
RL immunoglobulin epsilon chain expressed in a myeloma cell line.";  
RM Proc. Natl. Acad. Sci. U.S.A. 79:6661-6665(1982).  
[7]  
RP 3D-STRUCTURE MODELING.  
RA Medline=87089848; PubMed=3796618;  
RX Padlan E.A., Davies D.R.;  
BT "A model of the Fc of immunoglobulin E";  
RL Mol. Immunol. 23:1063-1075(1986).  
CC -1- SIMILARITY: Contains 4 immunoglobulin-like domains.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; L00022; AAB59424.1; ALT\_INIT.  
DR PIR; A22771; EHHU.  
DR PDB; 1IGE; 15-JUL-92.  
DR PDB; 1PP5; 30-JAN-02.  
DR PDB; 1G84; 16-MAY-01.  
DR PDB; 100V; 18-SEP-02.  
DR Genew; HGNC:5522; IGHE.  
DR MIM; 147180; --  
DR GO; GO:0003823; P:antigen binding; NAS.  
DR GO; GO:0006955; P:immune response; NAS.  
DR InterPro; IPR007110; Ig\_Like.  
DR InterPro; IPR003597; Ig\_C1.  
DR Pfam; PF00047; Ig; 4.  
DR SMART; SM00407; IG1; 4.  
DR PROSITE; PS00835; IG\_LIKE; 4.  
DR PROSITE; PS00290; IG\_MHC; 3.  
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
KW 3D-structure; Repeat.  
FT NON\_TER 1  
FT DOMAIN 6 103 IG-LIKE 1.  
FT DOMAIN 112 210 IG-LIKE 2.  
FT DOMAIN 214 318 IG-LIKE 3.  
FT DOMAIN 324 423 IG-LIKE 4.  
FT DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).  
FT DISULFID 15 105  
FT DISULFID 29 85  
FT DISULFID 121 121 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 135 193  
FT DISULFID 209 209 INTERCHAIN (WITH A HEAVY CHAIN).  
FT DISULFID 239 299  
FT DISULFID 345 405  
FT CARBOHYD 21 21 N-LINKED (GLCNAC. . .).  
FT CARBOHYD 49 49 N-LINKED (GLCNAC. . .).  
FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .).  
FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .).  
FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .).  
FT CARBOHYD 275 275 N-LINKED (GLCNAC. . .).  
FT VARIANT 359 359 W -> L (POSSIBLE POLYMORPHISM).  
FT STRAND 110 110 /FTid=VAR\_003885.  
FT STRAND 113 117  
FT HELIX 122 124  
FT TURN 125 125  
FT STRAND 130 141  
FT STRAND 146 150  
FT STRAND 155 155  
FT STRAND 159 161  
FT STRAND 165 166  
FT TURN 168 169

FT STRAND 172 181  
FT HELIX 182 186  
FT TURN 187 188  
FT STRAND 192 196  
FT TURN 198 199  
FT TURN 201 202  
FT STRAND 205 207  
FT STRAND 218 222  
FT TURN 226 229  
FT TURN 231 232  
FT STRAND 236 242  
FT STRAND 252 252  
FT STRAND 255 257  
FT TURN 258 259  
FT STRAND 260 260  
FT TURN 265 266  
FT STRAND 271 274  
FT TURN 275 276  
FT STRAND 277 285  
FT HELIX 288 293  
FT TURN 294 294  
FT STRAND 297 302  
FT TURN 304 305  
FT STRAND 310 314  
FT STRAND 322 322  
FT STRAND 325 329  
FT HELIX 334 337  
FT STRAND 340 348  
FT STRAND 350 351  
FT STRAND 356 360  
FT TURN 366 369  
FT STRAND 374 374  
FT STRAND 377 378  
FT TURN 380 381  
FT STRAND 384 385  
FT STRAND 387 393  
FT HELIX 394 398  
FT TURN 399 400  
FT STRAND 404 408  
FT STRAND 418 421  
SQ SEQUENCE 428 AA; 47019 MW; 25C4CA072AA558A0 CRC64;  
Query Match 55.2%; Score 80; DB 1; Length 428;  
Best Local Similarity 60.9%; Pred. NO. 0.00038;  
Matches 14; Conservative 3; Mismatches 6; Indels 0; Gaps 0;  
QY 2 GEGYQSYVDHPHPKPIVRSITK 24  
DB 294 GETYQCRVTHPLPALNRSTTK 316  
RESULT 4  
HB2D RAT STANDARD; PRT; 264 AA.  
AC P18211;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE RT1 class II histocompatibility antigen, D-1 beta chain precursor.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=Lewis familiaris; TISSUE=Bone marrow;  
RX MEDLINE=90356406; PubMed=2388838;  
RA Syha-Jedelhauser J., Reske K.;  
RT "Sequence of rat cDNA clone pLR beta 112 coding for the RT1.D beta I chain.";  
RL Nucleic Acids Res. 18:4598-4598(1990).  
CC -1- FUNCTION: Involved in the presentation of foreign antigens to the immune system.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; X53054; CAA37221.1; --  
 PIR; S10989; S10989.  
 HSP; P13760; 25EB.  
 InterPro; IPR007110; Ig-like.  
 InterPro; IPR003597; Ig cl.  
 InterPro; IPR003006; Ig MHC.  
 InterPro; IPR000353; MHC\_II\_beta.  
 Pfam; PF00047; Ig; 1.  
 Pfam; PF00969; MHC\_II\_beta; 1.  
 ProDom; PD000328; MHC\_II\_beta; 1.  
 SMART; SM00407; IG1; 1.  
 PROSITE; PS50835; IG LIKE; 1.  
 PROSITE; PS00290; IG MHC; 1.  
 MHC II; Transmembrane; Glycoprotein; Signal.  
 SIGNAL 1 26  
 CHAIN 27 264

RT1 CLASS II HISTOCOMPATIBILITY ANTIGEN,  
 D-1 BETA CHAIN.  
 DOMAIN 27 120  
 DOMAIN 121 215  
 DOMAIN 216 228  
 TRANSMEM 227 248  
 DOMAIN 249 264  
 DISULFID 42 106  
 DISULFID 144 200  
 CARBOHYD 46 46  
 SEQUENCE 264 AA; 29876 MW; 158F357355177DA1 CRC64;

Query Match 35.2%; Score 51; DB 1; Length 264;  
 Best Local Similarity 47.1%; Pred. No. 3.6;  
 Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Oy 2 GEGYQSRVDHPHPKPI 18  
 Db 195 GEVYTCQVEHSLSPV 211

## RESULT 5

HALF\_CHICK STANDARD; PRT; 345 AA.  
 ID HALF\_CHICK  
 AC P15979;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Class I histocompatibility antigen, F10 alpha chain precursor (B-F  
 DE histocompatibility F10 antigen) (B-F-beta-IV) (B12).  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OC NCBI\_TaxID=9031;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=B12; TISSUE=Liver;  
 RX MEDLINE=9903642; PubMed=3411149;  
 RA Guillemot F., Billault A., Pourquie O., Behar G., Chausse A.M.,  
 RA Zoorob R., Kreibich G., Auffray C.;  
 RT "A molecular map of the chicken major histocompatibility complex: the  
 RT class II beta genes are closely linked to the class I genes and the  
 RT nuclear organizer".  
 RL EMBO J. 7:2775-2785(1988).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=90316612; PubMed=2370087;  
 RX Kroemer G., Zoorob R., Auffray C.;

RT  
 RL Immunogenetics 31:405-409(1990).  
 CC -!- FUNCTION: Involved in the presentation of foreign antigens to the  
 CC immune system.  
 CC -!- SUBUNIT: Heterodimer of an alpha chain and a beta chain (beta-2-  
 CC microglobulin).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

EMBL; X12780; CAA31272.1; --  
 EMBL; M31012; AAA48947.1; --  
 PIR; A45846; HLCB84.  
 HSP; P13599; 3FRU.  
 InterPro; IPR007110; Ig-like.  
 InterPro; IPR003597; Ig cl.  
 InterPro; IPR003006; Ig MHC.  
 InterPro; IPR001039; MHC\_I.  
 Pfam; PF00047; Ig; 1.  
 Pfam; PF00129; MHC\_I; 1.  
 PRINTS; PR01638; MHCCLASSI.  
 ProDom; PD000050; MHC\_I; 1.  
 SMART; SM00407; IG1; 1.  
 PROSITE; PS50835; IG LIKE; 1.  
 PROSITE; PS00290; IG MHC; 1.  
 MHC I; Transmembrane; Glycoprotein; Signal.  
 SIGNAL 1 22  
 CHAIN 23 345

CLASS I HISTOCOMPATIBILITY ANTIGEN, F10  
 ALPHA CHAIN.  
 DOMAIN 23 110  
 DOMAIN 111 201  
 DOMAIN 202 292  
 DOMAIN 293 301  
 TRANSMEM 302 324  
 DOMAIN 325 345  
 DISULFID 121 183  
 DISULFID 121 277  
 CARBOHYD 59 59  
 CARBOHYD 107 107  
 CONFLICT 339 339  
 SEQUENCE 345 AA; 38246 MW; 430DCCF8091869A4 CRC64;

Query Match 33.8%; Score 49; DB 1; Length 345;  
 Best Local Similarity 50.0%; Pred. No. 9.2;  
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Oy 2 GEGYQSRVDHPHPKPI 17  
 Db 272 GDKYQCRVEHASLPQP 287

## RESULT 6

MURE FUSNN STANDARD; PRT; 485 AA.  
 ID MURE FUSNN  
 AC Q8R635;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE UDP-N-acetylmuramoyl-D-glutamate--2,6-diaminopimelate ligase  
 DE (EC 6.3.2.13) (UDP-N-acetylmuramyl-tripeptide synthetase) (Meso-  
 DE diaminopimelate-adding enzyme) (UDP-MurNAC-tripeptide synthetase).  
 GN MURE OR FN1225.  
 OS Fusobacterium nucleatum (subsp. nucleatum).  
 OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;  
 OC Fusobacterium.  
 OC NCBI\_TaxID=76856;  
 RN [1]  
 RP SEQUENCE FROM N.A.

```

RX STRAIN=ATCC 25586;
RX MEDLINE=2186394; PubMed=1189109;
RA Kapatral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasileva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Fongstein M., Kyripides N., Overbeek R.,
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
RT nucleatum strain ATCC 25586."
RL J. Bacteriol. 184:2005-2018 (2002).
CC -!- FUNCTION: Cell wall formation. Diaminopimelic acid adding enzyme
CC (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl-L-alanyl-D-
CC glutamate + meso-2,6-diaminoheptanediate = ADP + phosphate + UDP-
CC N-acetylmuramoyl-L-alanyl-D-gamma-glutamyl-meso-2,6-diamino-
CC heptanediate.
CC -!- PATHWAY: Peptidoglycan biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: Belongs to the murCDEF family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE010628; AAL95421.1; -.
CC HAVAP; MF_00208; -.
CC InterPro; IPR000713; Mur_ligase.
CC InterPro; IPR004101; Mur_ligase_C.
CC InterPro; IPR005761; MurE.
CC Pfam; PF01225; Mur_ligase; 1.
CC Pfam; PF02875; Mur_ligase_C; 1.
CC TIGRFAMs; TIGR01085; murE; 1.
KW Peptidoglycan synthesis; Cell wall; Cell division; Ligase;
KW ATP-binding; Complete proteome.
FT NP_BIND 108 114 ATP (POTENTIAL).
FT SEQUENCE 485 AA; 55484 MW; E6299E43P2B7FF19 CRC64;
SQ
Query Match 33.8%; Score 49; DB 1; Length 485;
Best Local Similarity 40.9%; Pred. No. 13;
Matches 9; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
Qy 1 CGEQYQSRVDHPHPKPKIVRSI 22
Db 331 CGQYKIVDYAHTPDALVNV 352
RESULT 7
KPC1_CABEL STANDARD; PRT; 704 AA.
AC P34722;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Protein kinase C-like 1 (EC 2.7.1.-) (PKC) (Tetradecanoyl phorbol
DE acetate resistant protein 1).
GN TPA-1 OR B0545.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS A AND B).
RC STRAIN=Bristol N2;
RX MEDLINE=95387388; PubMed=7659466;
RA Sano T., Tabuse Y., Nishiwaki K., Miwa J.;
RT "The tpa-1 gene of Caenorhabditis elegans encodes two proteins
RT similar to Ca(2+)-independent protein kinase Cs: evidence by complete
RT genomic and complementary DNA sequences of the tpa-1 gene."
RJ J. Mol. Biol. 251:477-485 (1995).
```

```

RN SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RA Rohlfing T.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP REVISIONS, AND ALTERNATIVE SPLICING.
RA Waterston R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 148-704 FROM N.A.
RX MEDLINE=89186920; PubMed=2538925;
RA Tabuse Y., Nishiwaki K., Miwa J.;
RT "Mutations in a protein kinase C homolog confer phorbol ester
RT resistance on Caenorhabditis elegans."
RL Science 243:1713-1716 (1989).
RN [5]
RP REVISIONS.
RA Miwa J.;
RL Submitted (MAR-1993) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PKC is activated by diacylglycerol which in turn
CC phosphorylates a range of cellular proteins. PKC also serves as
CC the receptor for phorbol esters, a class of tumor promoters.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=a;
CC IsoId=P34722-1; Sequence=Displayed;
CC Name=b;
CC IsoId=P34722-2; Sequence=VSP_004744;
CC -!- SIMILARITY: Contains 2 zinc-dependent phorbol-ester and DAG
CC binding domains.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. PKC
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D49525; BAA08470.1; -.
CC EMBL; D49525; BAA08471.1; -.
CC EMBL; AF078781; AAC26916.1; -.
CC EMBL; AF078781; AAC26917.1; -.
CC EMBL; D14815; BAA03556.1; -.
CC FIR; S60117; S60117.
CC FIR; T33400; T33400.
CC HSSP; P28867; IPTQ.
CC WormPep; B0545.1a; CEI7353.
CC WormPep; B0545.1b; CEI7354.
CC InterPro; IPR008973; C2_CaLB.
CC InterPro; IPR002219; DAG-PE-bind.
CC InterPro; IPR000961; Pkinase C.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR008271; Ser_thr_kinase.
CC InterPro; IPR002290; Ser_thr_kinase.
CC InterPro; IPR001245; Tyr_pkinase.
CC Pfam; PF00130; DAG-PE-bind; 2.
CC Pfam; PF00069; Pkinase; 1.
CC Pfam; PF00433; Pkinase C; 1.
CC PRINTS; PR00008; DAGPEDOMAIN.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00109; C1; 2.
CC SMART; SM00133; S_TK_X; 1.
CC SMART; SM00220; S_TKC; 1.
CC PROSITE; PS00479; DAG_PE_BIND_DOM_1; 2.
CC PROSITE; PS00081; DAG_PE_BIND_DOM_2; 2.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
```

KW Transferase; Serine/threonine-protein kinase; ATP-binding; Zinc;  
 FT Phorbol-ester binding; Repeat; Alternative splicing; Phosphorylation.  
 FT DOMAIN 166 215 PHORBOL-ESTER AND DAG BINDING 1.  
 FT DOMAIN 238 287 PHORBOL-ESTER AND DAG BINDING 2.  
 FT PROTEIN KINASE  
 FT NP\_BIND 375 634 ATP (BY SIMILARITY).  
 FT BINDING 381 389 ATP (BY SIMILARITY).  
 FT BINDING 404 404 ATP (BY SIMILARITY).  
 FT ACT\_SITE 499 499 BY SIMILARITY.  
 FT MOD\_RES 89 89 PHOSPHORYLATION (AUTO-) (POTENTIAL).  
 FT MOD\_RES 139 139 PHOSPHORYLATION (AUTO-) (POTENTIAL).  
 FT MOD\_RES 324 324 PHOSPHORYLATION (AUTO-) (POTENTIAL).  
 FT VARSPLIC 1 137 Missing (in isoform b).  
 FT /FTID=VSP\_004744.  
 SQ SEQUENCE 704 AA; 80298 MW; C44F2E25F59057E3 CRC64;  
 Query Match 33.4%; Score 48.5; DB 1; Length 704;  
 Best Local Similarity 37.9%; Pred.No. 23;  
 Matches 11; Conservative 2; Mismatches 11; Indels 5; Gaps 1;  
 QY 2 GEGYQSRVD-----HPHFKPPIVRSITK 25  
 DB 577 GEGEDFDSILNERPYPFKTISKAAC 605  
 RESULT 8  
 GC1\_MOUSE STANDARD; PRT; 324 AA.  
 AC P01868;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Ig gamma-1 chain C region secreted form.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 [1]  
 SEQUENCE FROM N.A.  
 RX MEDLINE=80045036; PubMed=115593;  
 RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,  
 RA Takahashi N., Mano Y.;  
 RT "Cloning and complete nucleotide sequence of mouse immunoglobulin  
 gamma 1 chain gene.";  
 RL Cell 48:559-568(1979).  
 RN [2]  
 RN SEQUENCE OF 76-324 FROM N.A. (MYELOMA PROTEIN MOPC 31C).  
 RP MEDLINE=80202559; PubMed=6769752;  
 RA Obata M., Yamawaki-Kataoka Y., Takahashi N., Kataoka T., Shimizu A.,  
 RA Mano Y., Seidman J.G., Peterlin B.M., Leder P., Honjo T.;  
 RT "Immunoglobulin gamma 1 heavy chain gene: structural gene sequences  
 cloned in a bacterial plasmid.";  
 RL Gene 9:87-97(1980).  
 RN [3]  
 RN SEQUENCE OF 70-322 FROM N.A. (MYELOMA PROTEIN MOPC 21).  
 RP MEDLINE=80012837; PubMed=113776;  
 RA Rogers J., Clarke P., Salsner W.;  
 RT "Sequence analysis of cloned cDNA encoding part of an immunoglobulin  
 heavy chain.";  
 RL Nucleic Acids Res. 6:3305-3321(1979).  
 RN [4]  
 RN SEQUENCE (MYELOMA PROTEIN MOPC 21).  
 RP MEDLINE=78242288; PubMed=98524;  
 RA Adetugbo K.;  
 RT "Evolution of immunoglobulin subclasses. Primary structure of a  
 murine myeloma gamma1 chain.";  
 RL J. Biol. Chem. 253:6068-6075(1978).  
 RN [5]  
 RN DISULFIDE BONDS (MOPC 21).  
 RP MEDLINE=73008889; PubMed=5073237;  
 RA Svasti J., Milstein C.;  
 RT "The disulphide bridges of a mouse immunoglobulin G1 protein.";  
 RL Biochem. J. 126:837-850(1972).  
 CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=2;  
 CC Name=Secreted;  
 CC IsoId=P01868-1; Sequence=Displayed;  
 CC Note=May be the major isoform;  
 CC Name=Membrane-bound;  
 CC IsoId=P01869-1; Sequence=External;  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; V00793; CAA24172.1; -;  
 CC EMBL; V00793; CAA24173.1; -;  
 CC EMBL; V00793; CAA24174.1; -;  
 CC EMBL; V00793; CAA24175.1; -;  
 CC EMBL; V00795; CAA24176.1; -;  
 CC PIR; A02159; G1MS.  
 CC GlycoSuiteDB; P01868; -;  
 CC MGD; MGI:96446; Igh-4  
 CC InterPro; IPR007110; Ig-like.  
 CC InterPro; IPR003597; Ig cl.  
 CC InterPro; IPR003006; Ig\_MHC.  
 CC Pfam; PF00047; Ig\_2.  
 CC SMART; SM00407; IgC1; 2.  
 CC PROSITE; PS50835; IG\_LIKE; 3.  
 CC PROSITE; PS00290; IG\_MHC; 1.  
 CC Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
 KW Alternative splicing.  
 KW NON\_TER 1  
 FT DOMAIN 1 97 CH1.  
 FT DOMAIN 98 110 HINGE.  
 FT DOMAIN 111 217 CH2.  
 FT DOMAIN 218 324 CH3.  
 FT DISULFID 27 82  
 FT DISULFID 102 102 INTERCHAIN (WITH A LIGHT CHAIN).  
 FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).  
 FT DISULFID 138 198  
 FT CARBOHYD 174 174 N-LINKED (GLCNAC...).  
 FT /FTID=CAR\_000055.  
 FT  
 FT DISULFID 244 302 REMOVED POST-TRANSLATIONALLY.  
 FT MOD\_RES 324 324 N -> D (IN REF. 3).  
 FT CONFLICT 276 276 N -> D (IN REF. 3).  
 FT CONFLICT 278 278 N -> D (IN REF. 3).  
 SQ SEQUENCE 324 AA; 35704 MW; A338812F3D1F2C93 CRC64;  
 Query Match 33.1%; Score 48; DB 1; Length 324;  
 Best Local Similarity 39.1%; Pred.No. 12;  
 Matches 9; Conservative 7; Mismatches 7; Indels 0; Gaps 0;  
 QY 2 GEGYQSRVDHPHFKPPIVRSITK 24  
 DB 193 GKEFKRVNSAFAFPAPKEKTISK 215  
 RESULT 9  
 GC1\_MOUSE STANDARD; PRT; 393 AA.  
 AC P01869;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Ig gamma-1 chain C region, membrane-bound form.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=80045036; PubMed=115593;
RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
RT Takahashi N., Mano Y.;
RA "Cloning and complete nucleotide sequence of mouse immunoglobulin
RT gamma 1 chain gene.";
RL Cell 18:559-568(1979).
RN [2]
RP SEQUENCE OF 323-393 FROM N.A.
RX MEDLINE=82197626; PubMed=6804950;
RA Tyler B.M., Cowan A.F., Gerondakis S.D., Adams J.M., Bernard O.;
RT "mRNA for surface immunoglobulin gamma chains encodes a highly
RT conserved transmembrane sequence and a 28-residue intracellular
RT domain.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2008-2012(1982).
RN [3]
RP SEQUENCE OF 323-366 FROM N.A.
RX MEDLINE=82115295; PubMed=6799207;
RA Rogers J., Choi E., Souza L., Carter C., Word C.J., Kuehl M.,
RT Eisenberg D., Wall R.;
RA "Gene segments encoding transmembrane carboxyl termini of
RT immunoglobulin gamma chains.";
RL Cell 26:19-27(1981).
RN [4]
RP SEQUENCE OF 1-44 FROM N.A.
RX MEDLINE=82222190; PubMed=6283537;
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
RT immunoglobulin gamma chains.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Membrane-bound;
CC IsoId=P01869-1; Sequence=Displayed;
CC Name=Secreted;
CC IsoId=P01868-1; Sequence=External;
CC Note=May be the major isoform;
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00793; CAA24172.1; -
DR EMBL; V00793; CAA24173.1; -
DR EMBL; V00793; CAA24174.1; -
DR PIR; B02159; GIMSM.
DR PDB; 15C8; 23-MAR-99.
DR PDB; 1A86; 18-MAR-98.
DR PDB; 1CL7; 12-JAN-00.
DR PDB; 1F11; 06-FEB-01.
DR PDB; 1F58; 29-DEC-99.
DR PDB; 1KCS; 24-JUL-02.
DR PDB; 1KCR; 11-MAY-02.
DR PDB; 25C8; 09-JUL-99.
DR MGD; MGI:96446; Igh-4.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig cl.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00407; IGC1; 2.
DR PROSITE; PS50835; IG LIKE; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing; Transmembrane; 3D-structure.
FT NON_TER 1 1
FT DOMAIN 98 97 CH1.
FT DOMAIN 98 110 HINGE.
FT DOMAIN 111 217 CH2.

```

```

FT DOMAIN 218 324 CH3.
FT DISULFID 27 82 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 138 198 N-LINKED (GLCNAC. . .).
FT CARBOHYD 174 174 POTENTIAL.
FT DISULFID 244 302 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 358 393 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 393 AA; 43386 MW; 4CC98343B7A1CE27 CRC64;

Query Match 33.1%; Score 48; DB 1; Length 393;
Best Local Similarity 39.1%; Pred. No. 15;
Matches 9; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPIVRSITK 24
Db 193 GKFKCRVNSAAPPAPIEKTIK 215

RESULT 10
GCA_RAT STANDARD; PRT; 322 AA.
ID_GCA_RAT
AC. P20760;
DI 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE IG gamma-2A chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
CC -1- SIMILARITY: Contains 3 immunoglobulin-like domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M13804; AAA41376.1; ALT_INIT.
DR PIR; PS0019; PS0019.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig cl.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00407; IGC1; 2.
DR PROSITE; PS50835; IG LIKE; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Repeat.
FT NON_TER 1 1
FT DOMAIN 6 98 IG-LIKE 1.
FT DOMAIN 115 212 IG-LIKE 2.
FT DOMAIN 221 317 IG-LIKE 3.
FT DISULFID 27 82 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 105 105 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 136 196
FT DISULFID 242 300
FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 322 AA; 35186 MW; B8EAL36A9DE01EDB CRC64;

```





```
CC AND AN OXIDOREDUCTASE, TO CATALYZE THE REDUCTION OF SUPEROXIDE TO
CC HYDROGEN PEROXIDE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Reduced rubredoxin + superoxide + 2 H(+) =
CC rubredoxin + H(2)O(2).
CC -1- COFACTOR: Iron (By similarity).
CC -1- SIMILARITY: Belongs to the desulfoferrodoxin family.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U67520; AAB98735.1; -.
CC F1R; E64392; E64392.
CC HSSP; P22076; 1DFX.
CC TIGR; M00741; -.
CC InterPro; IPR002742; Desulfoferrodox.
CC Pfam; PF01880; Desulfoferrodox; 1.
CC ProDom; PD006618; Desulfoferrodox; 1.
CC TIGRFAMs; TIGR00332; neela_ferrous; 1.
CC Hypothetical protein; Oxidoreductase; Electron transport; Iron;
KW Complete proteome.
FT METAL 20 20 IRON (BY SIMILARITY).
FT METAL 45 45 IRON (BY SIMILARITY).
FT METAL 52 52 IRON (BY SIMILARITY).
FT METAL 101 101 IRON (BY SIMILARITY).
FT METAL 104 104 IRON (BY SIMILARITY).
SQ SEQUENCE 116 AA; 13950 MW; 4CF2C76237DE0673 CRC64;
Query Match 31.7%; Score 46; DB 1; Length 116;
Best Local Similarity 43.5%; Pred. No. 8;
Matches 10; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
QY 2 GEGYQSRVDHPHPKPIVRSITX 24
D 62 GDLYLARVDTQPMKEVXLVX 84
RESULT 13
ID VNST CVBQ STANDARD; PRT; 277 AA.
AC P18517;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 32 kDa nonstructural protein.
OS Bovine coronavirus (strain Quebec) (BCoV) (BCV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=11133;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89345182; PubMed=2762160;
RA Cox G.J., Parker M.D., Babiuk L.A.;
RT "The sequence of cDNA of bovine coronavirus 32K nonstructural gene.";
RL Nucleic Acids Res. 17:5847-5847(1989).
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X15445; CAA33485.1; -.
CC F1R; A34039; MNH32.
CC InterPro; IPR007878; Corona_NS2A.
CC Pfam; PF05213; Corona_NS2A; 1.
CC Nonstructural protein.
KW
```

```
SQ SEQUENCE 277 AA; 31896 MW; D2EC7BAFB86EA11 CRC64;
Query Match 31.7%; Score 46; DB 1; Length 277;
Best Local Similarity 58.3%; Pred. No. 20;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 2 GEGYQSRVDHPH 13
D 78 GEGFIKFDNPH 89
RESULT 14
ID GCB_MOUSE STANDARD; PRT; 336 AA.
AC P01866;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE IG gamma-2B chain C region secreted form.
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ALLELE A).
RX MEDLINE=80120716; PubMed=5766534;
RA Yanawaki-Kataoka Y., Kataoka T., Takahashi N., Obata M., Honjo T.;
RT "Complete nucleotide sequence of immunoglobulin gamma2b chain gene
RT cloned from newborn mouse DNA.";
RL Nature 283:786-789(1980).
RN [2]
RP SEQUENCE FROM N.A. (MPC 11).
RX MEDLINE=80081501; PubMed=117548;
RA Tucker P.W., Marcu K.B., Slightom J.L., Blattner F.R.;
RT "Structure of the constant and 3' untranslated regions of the murine
RT gamma 2b heavy chain messenger RNA.";
RL Science 206:1299-1303(1979).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=80081502; PubMed=117549;
RA Tucker P.W., Marcu K.B., Newell N., Richards J., Blattner F.R.;
RT "Sequence of the cloned gene for the constant region of murine gamma
RT 2b immunoglobulin heavy chain.";
RL Science 206:1303-1306(1979).
RN [4]
RP SEQUENCE FROM N.A. (ALLELE B).
RX MEDLINE=82173203; PubMed=6803173;
RA Ollo R., Rougeon F.;
RT "Mouse immunoglobulin allotypes: post-duplication divergence of gamma
RT 2a and gamma 2b chain genes.";
RL Nature 296:761-763(1982).
RN [5]
RP CARBOHYDRATE-LINKAGE SITE THR-105.
RX MEDLINE=94216359; PubMed=7512967;
RA Kim H., Yamaguchi Y., Masuda K., Matsunaga C., Yamamoto K.,
RA Irimura T., Takahashi N., Kato K., Arata Y.;
RT "O-glycosylation in hinge region of mouse immunoglobulin G2b.";
RL J. Biol. Chem. 269:12345-12350(1994).
CC -1- SUBCELLULAR LOCATION: Secreted (Potential).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Secreted;
CC IsoId=P01866-1; Sequence=Displayed;
CC Note=May be the major isoform;
CC Name=Membrane-bound;
CC IsoId=P01867-1; Sequence=External;
CC -1- PTM: O-LINKED GLYCAN CONSISTS OF GAL-GALNAc DISACCHARIDE WHICH IS
CC MODIFIED WITH 2 SIALIC ACID RESIDUES.
CC -1- PTM: DISULFIDE BONDS BY HOMOLOG WITH OTHER IG GAMMA CHAINS.
CC -1- MISCELLANEOUS: The a allele sequence is shown.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like domains.
DR PIR; S25057; G2MS11.
DR HSSP; P01842; 7FAB.
```

```

DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig cl.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00407; IGC1; 2.
DR PROSITE; PS00835; IG_LIKE; 3.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing; Repeat.
FT NON_TER 1
FT DOMAIN 6 98 IG-LIKE 1.
FT DOMAIN 127 226 IG-LIKE 2.
FT DOMAIN 235 331 IG-LIKE 3.
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 118 118 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 150 210
FT DISULFID 255 314
FT CARBOHYD 105 105 O-LINKED (GALNAc...),
FT MOD_RES 336 336 REMOVED POST-TRANSLATIONALLY (PROBABLE).
FT VARIANT 163 163 Q -> R (IN ALLELE B).
FT VARIANT 194 194 T -> A (IN ALLELE B).
FT VARIANT 300 300 N -> D (IN ALLELE B).
FT VARIANT 301 301 M -> I (IN ALLELE B).
FT CONFLICT 25 25 L -> S (IN REF. 2 AND 3).
FT CONFLICT 36 36 S -> P (IN REF. 2 AND 3).
FT CONFLICT 239 239 I -> T (IN REF. 2 AND 3).
FT CONFLICT 239 239
FT SEQUENCE 336 AA; 36658 MW; 7D879662607C356E CRC64;

Query Match 31.7%; Score 46; DB 1; Length 336;
Best Local Similarity 34.8%; Pred. No. 24;
Matches 8; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPFKPIVRSITK 24
| : : : : | : : : : |
DB 205 GKFKCKVNNKDLPSPIERTISK 227

RESULT 15
ID _DCUP_SCHPO STANDARD; PRT; 370 AA.
AC Q9USJ5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Uroporphyrinogen decarboxylase (EC 4.1.1.37) (URO-D) (UPD).
GN SPCC4B3.05C
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Helroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,

```

---

```

RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Furnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aveq S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RA "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -!- CATALYTIC ACTIVITY: Uroporphyrinogen-III = coproporphyrinogen + 4
CO(2).
CC -!- PATHWAY: Porphyrin and heme biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the uroporphyrinogen decarboxylase family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; AL132870; CAB60679.1; -.
DR PIR; T50443; T50443.
DR HSP; P06132; 1URO.
DR GeneDB; Spombe; SPCC4B3.05c; -.
DR InterPro; IPR008361; Heme.
DR InterPro; IPR000257; Uro_decarbxylys.
DR Pfam; PF01208; URO-D; 1.
DR ProDom; PD003225; Uro_decarbxylys; 1.
DR TIGRFAMs; TIGR01464; heme; 1.
DR PROSITE; PS00906; UROD_1; FALSE_NEG.
DR PROSITE; PS00907; UROD_2; 1.
DR Lyase; Decarboxylase; Porphyrin biosynthesis; Heme biosynthesis.
KW SEQUENCE 370 AA; 41846 MW; 110DFC463733FD89 CRC64;
SQ

```

Query Match 31.7%; Score 46; DB 1; Length 370;  
Best Local Similarity 87.5%; Pred. No. 27;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 12 PHFKPIV 19  
| | | | | | | |  
DB 110 PHFKPLV 117

Search completed: February 26, 2004, 08:24:21  
Job time : 7.4 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 08:18:11 ; Search time 31.4 Seconds  
(without alignments)  
251.209 Million cell updates/sec

Title: US-09-701-623C-7

Perfect score: 145

Sequence: 1 CEGYQSRVDHPHFPKPIVRSITKC 25

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_25:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phase:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_rvirus:\*

16: sp\_bacteriaph:\*

17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62	42.8	661	5 Q86P90	Q86P90 drosophila
2	62	42.8	2594	5 Q9VM88	Q9VM88 drosophila
3	62	42.8	2922	5 Q8IFU3	Q8IFU3 drosophila
4	54	37.2	317	16 Q98DQ7	Q98DQ7 rhizobium 1
5	54	37.2	343	10 Q7XV01	Q7XV01 oryza sativ
6	54	37.2	648	5 Q9VJU4	Q9VJU4 drosophila
7	54	37.2	648	5 Q9KXD7	Q9KXD7 drosophila
8	54	37.2	701	5 Q86BL2	Q86BL2 drosophila
9	52	35.9	217	7 Q30827	Q30827 ovis aries
10	51	35.2	245	7 Q31270	Q31270 rattus norv
11	51	35.2	261	7 Q9TQA7	Q9TQA7 rattus norv
12	51	35.2	264	7 Q9TQA5	Q9TQA5 rattus norv
13	50	34.5	140	12 Q98A59	Q98A59 bovine vira
14	50	34.5	142	7 Q95336	Q95336 pongo pygma
15	50	34.5	314	11 Q9ESH2	Q9ESH2 rattus norv
16	50	34.5	337	6 Q95M34	Q95M34 equus cabal

#### ALIGNMENTS

##### RESULT 1

Q86P90 ID Q86P90 PRELIMINARY; PRT; 661 AA.

AC Q86P90; DT 01-JUN-2003 (TREMBLrel. 24, Created)

DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)

DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE GH08772p.

GN CG11321.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7227;

RN [1]\_TaxID=7227;

RP SEQUENCE FROM N.A.

RC STRAIN=;

RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,

RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,

RA George R., Gonzalez C., Guarin H., Kronmiller B., Li P., Liao G.,

RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,

RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,

RA Celisner S.,

RL Submitted (JAN-2003) to the ENBL/GenBank/DBJ databases.

DR EMBL; BT003307; AAC05067.1; -

DR InterPro; IPR001876; Znf.RangDP.

DR SMART; SM00547; Znf.RBZ; 1.

DR PROSITE; PS01358; ZF.RANBP2.1; 1.

SQ SEQUENCE 661 AA; 72438 MW; FAE8471548E67FA4 CRC64;

Query Match 42.8%; Score 62; DB 5; Length 661;

Best Local Similarity 52.9%; Pred. No. 0.54;

Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 CEGYQSRVDHPHFPKPI 18

DB 211 CEGYMSQQHHPYPPV 227

17	50	34.5	357	10	Q947H5	Q947H5 nicotiana t
18	50	34.5	357	10	Q947H4	Q947H4 nicotiana t
19	50	34.5	493	2	Q84ON7	Q84ON7 streptococ
20	50	34.5	539	5	Q965D2	Q965D2 trypanosoma
21	49	33.8	87	9	Q8SCC4	Q8SCC4 vibrio harv
22	49	33.8	90	7	Q7YPM5	Q7YPM5 gallus gall
23	49	33.8	90	7	Q7YPM7	Q7YPM7 gallus gall
24	49	33.8	90	7	Q7YPM3	Q7YPM3 gallus gall
25	49	33.8	125	11	Q9CZ59	Q9CZ59 mus musculu
26	49	33.8	173	7	Q95594	Q95594 gallus gall
27	49	33.8	273	7	Q98194	Q98194 acrocephalu
28	49	33.8	338	7	Q31412	Q31412 gallus gall
29	49	33.8	338	7	Q8SPD0	Q8SPD0 gallus gall
30	49	33.8	344	7	Q46792	Q46792 gallus gall
31	49	33.8	351	7	Q95593	Q95593 gallus gall
32	49	33.8	355	7	Q46788	Q46788 gallus gall
33	49	33.8	355	7	Q8SPC8	Q8SPC8 gallus gall
34	49	33.8	355	7	Q8SPC7	Q8SPC7 gallus gall
35	49	33.8	355	7	Q46790	Q46790 gallus gall
36	49	33.8	355	7	Q8SPD1	Q8SPD1 gallus gall
37	49	33.8	355	7	Q8SPC9	Q8SPC9 gallus gall
38	49	33.8	355	7	Q46791	Q46791 gallus gall
39	49	33.8	355	7	Q46789	Q46789 gallus gall
40	49	33.8	355	7	Q86010	Q86010 gallus gall
41	49	33.8	355	7	Q95601	Q95601 gallus gall
42	49	33.8	355	7	Q9GIP6	Q9GIP6 gallus gall
43	49	33.8	355	7	Q31400	Q31400 gallus gall
44	49	33.8	359	7	Q98193	Q98193 acrocephalu
45	49	33.8	359	7	Q98192	Q98192 acrocephalu

RESULT 2  
Q9VM88  
ID Q9VM88 PRELIMINARY; PRT; 2594 AA.  
AC Q9VM88;  
DT 01-MAY-2000 (TRENBLrel. 13, Created)  
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)  
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
DE CG11321 protein.  
GN CG11321.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_taxid=7227;  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkely;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celnikier S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA Amanatides P.G., Scherer S.E., Richards S., Ashburner M., Henderson S.N.,  
RA George R.A., Lewis S.E., Yandell M.D., Zhang Q., Chen L.X.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champagne M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.N., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Deicher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Haris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.B., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasok P., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*."  
RL Science 287:2185-2195 (2000).  
[2]  
RP SEQUENCE FROM N.A.  
RA Celnikier S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,  
RA Carlson J.W., Center A., Champagne M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Fafian D.,  
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,  
RA Gonzales M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattel B., Moshrefi A.,  
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,  
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,  
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
RT "Sequencing of *Drosophila melanogaster* genome."  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

[3]  
RP SEQUENCE FROM N.A.  
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,  
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celnikier S.E.,  
RA Clump M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,  
RA Seale S.M.J., Smith E., Shu S., Smutniak F., Whitfield E., S.E.;  
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
RT "Annotation of *Drosophila melanogaster* genome."  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
[4]  
RP SEQUENCE FROM N.A.  
RA Adams M.D., Celnikier S.E., Gibbs R.A., Rubin G.M., Venter C.J.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
[5]  
RP SEQUENCE FROM N.A.  
RA FlyBase;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE003615; AAF52435.2; -;  
DR FlyBase; FBGN0031857; CG11321.  
DR InterPro; IPR002867; Znf C6HC.  
DR InterPro; IPR001876; Znf RangDP.  
DR InterPro; IPR001841; Znf\_ring.  
DR Pfam; PF01485; IER; 1.  
DR SMART; SM00547; ZNF RBZ; 1.  
DR PROSITE; PS01358; ZF\_RANBP2\_1; 1.  
DR PROSITE; PS00089; ZF\_RING\_2; 1.  
SQ SEQUENCE 2594 AA; 288532 MW; A9A23E25417E9447 CRC64;  
Query Match 42.8%; Score 62; DB 5; Length 2594;  
Best Local Similarity 52.9%; Pred. No. 2.2;  
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
QY 2 GEGYQSRVDHPHPKPI 18  
DB 384 GEGYMSQQHHPYPPV 400  
RESULT 3  
Q8IPJ3  
ID Q8IPJ3 PRELIMINARY; PRT; 2922 AA.  
AC Q8IPJ3;  
DT 01-MAR-2003 (TRENBLrel. 23, Created)  
DT 01-OCT-2003 (TRENBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE CG11321-PB.  
GN CG11321.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_taxid=7227;  
[1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champagne M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Deicher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Haris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.B., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasok P., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*."  
RL Science 287:2185-2195 (2000).  
[2]  
RP SEQUENCE FROM N.A.  
RA Celnikier S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,  
RA Carlson J.W., Center A., Champagne M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Fafian D.,  
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,  
RA Gonzales M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattel B., Moshrefi A.,  
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,  
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,  
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
RT "Sequencing of *Drosophila melanogaster* genome."  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Mishina N.V., Mobarri C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,  
RA Williams S.M., Woodager, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhou X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*.";  
RL Science 287:2185-2195(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Ceiniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,  
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dreenek D., Farfan D.,  
RA Ferreira S., Frise E., Galle R.E., Garg N.S., George R.A.,  
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
RA Ibegwam S.C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,  
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunco J.,  
RA Pacieb J.S., Paragas V., Park S., Patel S., Pfeiffer B.,  
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
RA Stapleton M., Strong R., Swirskas R., Tector C., Tyler D.,  
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
RT "Sequencing of *Drosophila melanogaster* genome.";  
RN Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,  
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Ceiniker S.E.,  
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,  
RA Seale S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,  
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
RT "Annotation of *Drosophila melanogaster* genome.";  
RN Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Adams M.D., Ceiniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RA FlyBase;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE003615; AAN10598.1; .  
DR FlyBase; FBgn0031857; CG11321.  
DR InterPro; IPR002867; Znf\_C6HC.  
DR InterPro; IPR001876; Znf\_RanGDP.  
DR InterPro; IPR001841; Znf\_ring.  
DR Pfam; PF01495; IBR; 1.  
DR SMART; SM00647; IBR; 1.  
DR SMART; SM00184; RING; 2.  
DR SMART; SM00547; Znf\_RBZ; 1.  
DR PROSITE; PS01358; ZF\_RANBP2\_1; 1.  
DR PROSITE; PS00089; ZF\_RING\_2; 1.  
SQ SEQUENCE 2922 AA; 324447 MW; FFABF47B0045540D CRC64;

Query Match 42.88; Score 62; DB 5; Length 2922;  
Best Local Similarity 52.99; Pred. No. 2.5; Indels 0; Gaps 0;  
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPI 18  
DB 437 GPGYMSQQHPPYPPV 453  
RESULT 4  
ID Q98DQ7 PRELIMINARY; PRT; 317 AA.  
AC Q98DQ7;  
DT 01-OCT-2001 (TRENBLrel. 18, Created)  
DT 01-OCT-2001 (TRENBLrel. 18, Last sequence update)  
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)  
DE Hypothetical protein mlr4594.  
GN MLR4594.  
OS Rhizobium loti (Mesorhizobium loti).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Phyllobacteriaceae; Mesorhizobium.  
OX NCBI\_TaxID=381;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MAFF303099;  
RX MEDLINE=21082930; PubMed=11214968;  
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
RA Watanabe A., Idesawa K., Ishikawa K., Kawashima K., Kimura T.,  
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,  
RA Takeuchi C., Yamada M., Tabata S.,  
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
RT Mesorhizobium loti.";  
RL DNA Res. 7:331-338(2000).  
DR EMBL; AP003004; BAB51213.1; .  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 317 AA; 33121 MW; F7DE20BB4F661F32 CRC64;  
Query Match 37.28; Score 54; DB 16; Length 317;  
Best Local Similarity 38.58; Pred. No. 4.4; Indels 4; Gaps 1;  
Matches 10; Conservative 6; Mismatches 6; Indels 4; Gaps 1;  
QY 1 CGEGYQSRVDH---PHPPKPIVRSI 22  
DB 256 CGEGYQSRVSVSPDPSTPTPIVYSL 281  
RESULT 5  
ID Q7XV01 PRELIMINARY; PRT; 343 AA.  
AC Q7XV01;  
DT 01-OCT-2003 (TRENBLrel. 25, Created)  
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE OSUNBA0036B21.18 protein.  
GN OSUNBA0036B21.18.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,  
RA Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Weng Q.J., Zhang L.,  
RA Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T.,  
RA Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.,  
RA Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,  
RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,  
RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,  
RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,  
RA Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL606636; CAD40900.1; .  
SQ SEQUENCE 343 AA; 36998 MW; BE4712E39FCEA175 CRC64;

Query Match 37.2%; Score 54; DB 10; Length 343;  
Best Local Similarity 50.0%; Pred. No. 4.8;  
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 EGYGSRVDHPHFKPIVR 20  
DB 311 QGFDDAVNPPSVIVR 328

RESULT 6  
Q9VJTU4 ID Q9VJTU4 PRELIMINARY; PRT; 648 AA.  
AC Q9VJTU4;  
DT 01-MAY-2000 (T-EMBLrel. 13, Created)  
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)  
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)  
DE CG18146 protein.  
GN BG:D800180.14 OR CG18146.  
OS Drosophila melanogaster (fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
ON NCBI\_TaxID=7227;  
RX STRAIN=Berkely;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkely;  
RL MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballwee R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Bertram M.R., Bouck J.J., Brokstein P., Brotter P.,  
RA Borkova D., Botchan M.R., Bouck J.J., Brokstein P., Brotter P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cavley S., Dahlike C., Davenport L.B., Davies P.,  
RA de Pablo S., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang X., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Mount S.M., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleib J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J.R., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skuseki M.P., Smith T.,  
RA Spier E., Spradling A.C., Stappleton M., Strong R., Sun E.,  
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.Q.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zhong X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster."  
RL Science 287:2185-2195 (2000).

SEQUENCE FROM N.A.

Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,  
Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
Ferrera S., Frise E., Galle R.F., Garg N.S., George R.A.,  
Holt R.A., Krommiller B., Manatides P.G., Brandon R.C., Rogers Y.,  
Banzone J., An H., Baldwin D., Banzone J., Beeson K.Y., Busam D.A.,  
Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
Ferrera S., Frise E., Galle R.F., Garg N.S., George R.A.,  
Holt R.A., Krommiller B., Manatides P.G., Brandon R.C., Rogers Y.,  
Banzone J., An H., Baldwin D., Banzone J., Beeson K.Y., Busam D.A.,  
Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
Ferrera S., Frise E., Galle R.F., Garg N.S., George R.A.,  
Holt R.A., Krommiller B., Manatides P.G., Brandon R.C., Rogers Y.,  
Banzone J., An H., Baldwin D., Banzone J., Beeson K.Y., Busam D.A.,  
Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
Ferrera S., Frise E., Galle R.F., Garg N.S., George R.A.,  
Holt R.A., Krommiller B., Manatides P.G., Brandon R.C., Rogers Y.,  
Banzone J., An H., Baldwin D., Banzone J., Beeson K.Y., Busam D.A.,  
Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
Ferrera S., Frise E., Galle R.F., Garg N.S., George R.A.,  
Holt R.A., Krommiller B., Manatides P.G., Brandon R.C., Rogers Y.,  
Banzone J., An H., Baldwin D., Banzone J., Beeson K.Y., Busam D.A.,  
Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
Ferrera S., Frise E., Galle R.F., Garg N.S., George R.A.,  
Holt R.A., Krommiller B., Manatides P.G., Brandon R.C., Rogers Y.,  
Banzone J., An H., Baldwin D., Banzone J., Beeson K.Y., Busam D.A.,  
Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
Ferrera S., Frise E., Galle R.F., Garg N.S., George R.A.,  
Holt R.A., Krommiller B., Manatides P.G., Brandon R.C., Rogers Y.,  
Banzone J., An H., Baldwin D., Banzone J., Beeson K.Y., Busam D.A.,  
Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
Ferrera S., Frise E., Galle R.F., Garg N.S., George R.A.,  
Holt R.A., Krommiller B., Manatides P.G., Brandon R.C., Rogers Y.,  
Banzone J., An H., Baldwin D., Banzone J., Beeson K.Y., Busam D.A.,  
Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
Ferrera S., Frise E., Galle R.F., Garg N.S., George R.A.,  
Holt R.A., Krommiller B., Manatides P.G., Brandon R.C., Rogers Y.,  
Banzone J., An H., Baldwin D., Banzone J., Beeson K.Y., Busam D.A.,  
Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
Ferrera S., Frise E., Galle R.F., Garg N.S., George R.A.,  
Holt R.A., Krommiller B., Manatides P.G., Brandon R.C., Rogers Y.,  
Banzone J., An H., Baldwin D., Banzone J., Beeson K.Y., Busam D.A.,  
Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
Ferrera S., Frise E., Galle R.F., Garg N.S., George R.A.,  
Holt R.A., Krommiller B., Manatides P.G., Brandon R.C., Rogers Y.,  
Banzone J., An H., Baldwin D., Banzone J., Beeson K.Y., Busam D.A.,  
Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
Ferrera S., Frise E., Galle R.F., Garg N.S., George R.A.,  
Holt R.A., Krommiller B., Manatides P.G., Brandon R.C., Rogers Y.,  
Banzone J., An H., Baldwin D., Banzone J., Beeson K.Y., Busam D.A.,  
Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
Ferrera S., Frise E., Galle R.F., Garg N.S., George R.A.,  
Holt R.A., Krommiller B., Manatides P.G., Brandon R.C., Rogers Y.,  
Banzone J., An H., Baldwin D., Banzone J., Beeson K.Y., Busam D.A.,  
Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
Ferrera S., Frise E., Galle R.F., Garg N.S., George R.A.,  
Holt R.A., Krommiller B., Manatides P.G., Brandon R.C., Rogers Y.,  
Banzone J., An H., Baldwin D., Banzone J., Beeson K.Y., Busam D.A.,  
Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
Ferrera S., Frise E., Galle R.F., Garg N.S., George R.A.,  
Holt R.A., Krommiller B., Manatides P.G., Brandon R.C., Rogers Y.,  
Banzone J., An H., Baldwin D., Banzone J., Beeson K.Y., Busam D.A.,  
Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
Ferrera S., Frise E., Galle R.F., Garg N.S., George R.A.,  
Holt R.A., Krommiller B., Manatides P.G., Brandon R.C., Rogers Y.,  
Banzone J., An H., Baldwin D., Banzone J., Beeson K.Y., Busam D.A.,  
Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
Ferrera S., Frise E., Galle R.F., Garg N.S., George R.A.,  
Holt R.A., Krommiller B., Manatides P.G., Brandon R.C., Rogers Y.,  
Banzone J., An H., Baldwin D., Banzone J., Beeson K.Y., Busam D.A.,  
Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
Ferrera S., Frise E., Galle R.F., Garg N.S., George R.A.,  
Holt R.A., Krommiller B., Manatides P.G., Brandon R.C., Rogers Y.,  
Banzone J., An H., Baldwin D., Banzone J., Beeson K.Y., Busam D.A.,  
Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
Ferrera S., Frise E., Galle R.F., Garg N.S., George R.A.,  
Holt R.A., Krommiller B., Manatides P.G., Brandon R.C., Rogers Y.,  
Banzone J., An H., Baldwin D., Banzone J., Beeson K.Y., Busam D.A.,  
Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek

Query Match 37.2%; Score 54; DB 10; Length 343;  
Best Local Similarity 50.0%; Pred. No. 4.8;  
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 EGYGSRVDHPHFKPIVR 20  
||| ||| |||  
DB 311 QGFDDAVNPPSVIVR 328

RESULT 6  
Q9VJTU4 PRELIMINARY; PRT; 648 AA.

ID OSVTJ4 AC OSVTJ4;  
DT 01-MAY-2000 (T-EMBLrel. 13, Created)  
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)  
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)  
DE CG18146 protein.  
GN BG:D800180.14 OR CG18146.  
OS Drosophila melanogaster (fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
ON NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkeley;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballwee R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Beriman M.R., Bouck J.J., Brokstein P., Brotter P.,  
RA Borkova D., Botchan M.R., Bouck J.J., Brokstein P., Butler A., Chandra I.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cavley S., Dahlike C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang X., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Moulton G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleib J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J.R., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skuseki M.P., Smith T.,  
RA Spier E., Spradling A.C., Stappleton M., Strong R., Sun E.,  
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.Q.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zhong X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster."  
RL Science 287:2185-2195 (2000).

SEQUENCE FROM N.A.  
CElniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
RA Banon J., An H., Baldwin D., Bonzon J., Beeson K.Y., Busam D.A.,  
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
RA Ferreira S., Frise E., Galie R.F., Garg N.S., George R.A.,  
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
RA Banon J., An H., Baldwin D., Bonzon J., Beeson K.Y., Busam D.A.,  
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
RA Ferreira S., Frise E., Galie R.F., Garg N.S., George R.A.,  
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
RA Banon J., An H., Baldwin D., Bonzon J., Beeson K.Y., Busam D.A.,  
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
RA Ferreira S., Frise E., Galie R.F., Garg N.S., George R.A.,  
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
RA Banon J., An H., Baldwin D., Bonzon J., Beeson K.Y., Busam D.A.,  
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
RA Ferreira S., Frise E., Galie R.F., Garg N.S., George R.A.,  
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
RA Banon J., An H., Baldwin D., Bonzon J., Beeson K.Y., Busam D.A.,  
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
RA Ferreira S., Frise E., Galie R.F., Garg N.S., George R.A.,  
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
RA Banon J., An H., Baldwin D., Bonzon J., Beeson K.Y., Busam D.A.,  
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
RA Ferreira S., Frise E., Galie R.F., Garg N.S., George R.A.,  
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
RA Banon J., An H., Baldwin D., Bonzon J., Beeson K.Y., Busam D.A.,  
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
RA Ferreira S., Frise E., Galie R.F., Garg N.S., George R.A.,  
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
RA Banon J., An H., Baldwin D., Bonzon J., Beeson K.Y., Busam D.A.,  
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
RA Ferreira S., Frise E., Galie R.F., Garg N.S., George R.A.,  
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
RA Banon J., An H., Baldwin D., Bonzon J., Beeson K.Y., Busam D.A.,  
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
RA Ferreira S., Frise E., Galie R.F., Garg N.S., George R.A.,  
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
RA Banon J., An H., Baldwin D., Bonzon J., Beeson K.Y., Busam D.A.,  
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
RA Ferreira S., Frise E., Galie R.F., Garg N.S., George R.A.,  
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
RA Banon J., An H., Baldwin D., Bonzon J., Beeson K.Y., Busam D.A.,  
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
RA Ferreira S., Frise E., Galie R.F., Garg N.S., George R.A.,  
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
RA Banon J., An H., Baldwin D., Bonzon J., Beeson K.Y., Busam D.A.,  
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
RA Ferreira S., Frise E., Galie R.F., Garg N.S., George R.A.,  
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
RA Banon J., An H., Baldwin D., Bonzon J., Beeson K.Y., Busam D.A.,  
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
RA Ferreira S., Frise E., Galie R.F., Garg N.S., George R.A.,  
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
RA Banon J., An H., Baldwin D., Bonzon J., Beeson K.Y., Busam D.A.,  
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
RA Ferreira S., Frise E., Galie R.F., Garg N.S., George R.A.,  
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
RA Banon J., An H., Baldwin D., Bonzon J., Beeson K.Y., Busam D.A.,  
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
RA Ferreira S., Frise E., Galie R.F., Garg N.S., George R.A.,  
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
RA Banon J., An H., Baldwin D., Bonzon J., Beeson K.Y., Busam D.A.,  
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
RA Ferreira S., Frise E., Galie R.F., Garg N.S., George R.A.,  
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D

```

RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Celnik S.E., Agbayani A., Arcaina T.T., Baxter E., Blazej R.G.,
RA Butenhoff C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
RA Houston K.A., Hummasti S.R., Karra K., Kearney L., Kim E., Lee B.,
RA Lewis S., Li P., Lonotan M.A., Mazda P., Moshrefi A.K., Moshrefi M.,
RA Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequiera A.,
RA Sethi H., Snir E., Svirska R.R., Wan K.H., Weinburg T., Zhang R.,
RA Zieran L.L., Rubin G.M.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003408; AAP44845.1; -
DR FlyBase; FBGN0028939; BG:DS00180.14.
DR InterPro; IPR003341; DUF139.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR Pfam; PF02363; C tripleX; 16.
DR SMART; SM00181; EGF; 10.
DR PROSITE; PS01186; EGF 2; 6.
KW Hypothetical protein_2.
SQ SEQUENCE 648 AA; 69419 MW; 3F16E0BFF4B94CF8 CRC64;

Query Match 37.2%; Score 54; DB 5; Length 648;
Best Local Similarity 69.2%; Pred. No. 9.2;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPH 13
DB 45 CGKGYVSRKDRGH 57

RESULT 8
ID Q86BL2 PRELIMINARY; PRT; 701 AA.
AC Q86BL2
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG18146-PB.
GN BG:DS00180.14 OR CG18146.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA April J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.M., Benos P.V., Berman B.P., Shandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Buzan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cleyer S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fowler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

```

```

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodager, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2165-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
RA Tupy J.L., Bergman C.M., Berman B.P., Carlson J.W., Celnik S.E.,
RA Clump M.E., Drysdale R.A., Emmert D., Frise E., de Grey A.D.N.J.,
RA Harris N.L., Kronmiller B., Marshall B., Millburn G.H., Richter J.,
RA Russo S., Searle S.M.J., Smith E., Shu S., Smutniak P.,
RA Whitfield E.J., Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J.,
RA Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003642; AA041187.1; -
DR InterPro; IPR003341; DUF139.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR Pfam; PF02363; C tripleX; 16.
DR Pfam; PF00008; EGF; 1.
DR SMART; SM00181; EGF; 15.
DR PROSITE; PS01186; EGF 2; 6.
SQ SEQUENCE 701 AA; 75366 MW; 95159C04C9C09BD4 CRC64;

Query Match 37.2%; Score 54; DB 5; Length 701;
Best Local Similarity 69.2%; Pred. No. 10;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPH 13
DB 98 CGKGYVSRKDRGH 110

RESULT 9
ID Q30827 PRELIMINARY; PRT; 217 AA.
AC Q30827
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ovar-DOB protein (Fragment).
GN OVAR-DOB.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Finnish Landrace; Tissue=Blood;
RX MEDLINE=96128257; PubMed=8537127;
RA Wright H., Redmond J., Ballingall K.T.;
RT "The sheep orthologue of the HLA-DOB gene.";

```



```

RL Immunogenetics 43:76-79 (1996).
DR EMBL; Z49879; CAA90036.1; -.
DR EMBL; Z49880; CAA90036.1; JOINED.
DR HSSP; P13760; 2SEB.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0045012; P:MHC class II receptor activity; IEA.
DR GO; GO:0019884; P:antigen presentation, exogenous antigen; IEA.
DR GO; GO:0019886; P:antigen processing, exogenous antigen via M. . .; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig cl.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR00353; MHC_II_beta.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00969; MHC_II_beta; 1.
DR ProDom; PD000328; MHC_II_beta; 1.
DR SMART; SM00407; IGcl; 1.
DR PROSITE; PS0835; IG LIKE; 1.
DR PROSITE; PS0290; IG LIKE; 1.
DR Glycoprotein; MHC II; Transmembrane.
KW NON_TER 217
FT SEQUENCE 217 AA; 24707 MW; 1E246E7DBC2AC6B2 CRC64;

Query Match 35.9%; Score 52; DB 7; Length 217;
Best Local Similarity 52.9%; Pred. No. 6.1;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHHPFKPI 18
DB 194 GEVYTCVLDHPSLPSPV 210

RESULT 10
Q31270
ID Q31270 PRELIMINARY; PRT; 245 AA.
AC Q31270;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Rat MHC RT1 class II E-beta chain (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_taxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8606895; PubMed=3934274;
RA Robertson K.A.; McMaster W.R.;
RT "Complete structure of a rat RT1 E-beta chain: Extensive conservation
of MHC class II beta chains."
RL J. Immunol. 135:4095-4099 (1985).
DR EMBL; M12382; AAA41610.1; -.
DR PIR; I55951; I55951.
DR HSSP; P13760; 2SEB.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0045012; P:MHC class II receptor activity; IEA.
DR GO; GO:0019884; P:antigen presentation, exogenous antigen; IEA.
DR GO; GO:0019886; P:antigen processing, exogenous antigen via M. . .; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig cl.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR00353; MHC_II_beta.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00969; MHC_II_beta; 1.
DR ProDom; PD000328; MHC_II_beta; 1.
DR SMART; SM00407; IGcl; 1.
DR PROSITE; PS0835; IG LIKE; 1.
DR PROSITE; PS0290; IG LIKE; 1.
DR Glycoprotein; MHC II; Transmembrane.
KW NON_TER 1
FT SEQUENCE 245 AA; 28069 MW; 6B21DADB2A4A299D CRC64;

```

```

Query Match 35.2%; Score 51; DB 7; Length 245;
Best Local Similarity 47.1%; Pred. No. 9.8;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHHPFKPI 18
DB 176 GEVYTCVLDHPSLPSPV 192

RESULT 11
Q9TQA7
ID Q9TQA7 PRELIMINARY; PRT; 261 AA.
AC Q9TQA7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Class II MHC RT1.D(a) beta chain precursor.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_taxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ACI;
RX MEDLINE=99299366; PubMed=10369938;
RA Tian L., Wang M., Yu J., Kahan B.D., Stepkowski S.M.;
RT "Nucleotide sequences of three distinct complementary DNA clones
encoding rat class II major histocompatibility complex RT1.D beta-
chain proteins."
RL Immunogenetics 49:735-737 (1999).
DR EMBL; AF084932; AAD39082.1; -.
DR HSSP; P13760; 2SEB.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0045012; P:MHC class II receptor activity; IEA.
DR GO; GO:0019884; P:antigen presentation, exogenous antigen; IEA.
DR GO; GO:0019886; P:antigen processing, exogenous antigen via M. . .; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig cl.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR00353; MHC_II_beta.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00969; MHC_II_beta; 1.
DR ProDom; PD000328; MHC_II_beta; 1.
DR SMART; SM00407; IGcl; 1.
DR PROSITE; PS0835; IG LIKE; 1.
DR PROSITE; PS0290; IG LIKE; 1.
DR Glycoprotein; MHC II; Signal; Transmembrane.
KW SIGNAL 1 26
FT CHAIN 27 261 CLASS II MHC RT1.D(A) BETA CHAIN.
RP SEQUENCE 261 AA; 29496 MW; 48E220A69B374B3D CRC64;

Query Match 35.2%; Score 51; DB 7; Length 261;
Best Local Similarity 47.1%; Pred. No. 10;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHHPFKPI 18
DB 192 GEVYTCVLDHPSLPSPV 208

RESULT 12
Q9TQA5
ID Q9TQA5 PRELIMINARY; PRT; 264 AA.
AC Q9TQA5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Class II MHC RT1.D(u) beta chain precursor.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Wistar Furth;  
 RX MEDLINE=99299366; PubMed=10369938;  
 RA Tian L., Mang M., Yu J., Kahan B.D., Stepkowski S.M.;  
 RT "Nucleotide sequences of three distinct complementary DNA clones  
 RT encoding rat class II major histocompatibility complex RT1.D beta-  
 RT chain proteins.";  
 RL Immunogenetics 49:735-737(1999).  
 DR EMBL; AF084934; AAD39084.1; -;  
 DR PIR; A60497; A60497.  
 DR HSSP; P13760; 2SER.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0045012; F:MHC class II receptor activity; IEA.  
 DR GO; GO:0019884; P:antigen presentation, exogenous antigen; IEA.  
 DR GO; GO:0019886; P:antigen processing, exogenous antigen via M. .; IEA.  
 DR GO; GO:0006955; P:immune response; IEA.  
 DR InterPro; IPR007110; IG-like.  
 DR InterPro; IPR003597; IG-cl.  
 DR InterPro; IPR003006; IG\_MHC.  
 DR Pfam; PF00047; IG; 1.  
 DR Pfam; PF00969; MHC II\_beta; 1.  
 DR ProDom; PD000328; MHC II\_beta; 1.  
 DR SMART; SM00407; IGcl; 1.  
 DR PROSITE; PS0835; IG\_LIKE; 1.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 KW Glycoprotein; MHC II; Signal; Transmembrane.  
 FT SIGNAL 1 26  
 FT CHAIN 27 264 CLASS II MHC RT1.D(U) BETA CHAIN.  
 SQ SEQUENCE 264 AA; 30138 MW; 3C888533514F531F CRC64;  
 Query Match 35.2%; Score 51; DB 7; Length 264;  
 Best Local Similarity 47.1%; Pred. No. 11;  
 Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;  
 QY 2 GEGYQSRVDHPHPFKPI 18  
 Db 195 GEGYTCQVEHPSPV 211  
 RESULT 13  
 Q99A59 PRELIMINARY; PRT; 140 AA.  
 AC Q99A59;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE E2 Glycoprotein (Fragment).  
 GN E2.  
 OS Bovine viral diarrhea virus strain 4998/89.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Pestivirus.  
 OX NCBI\_TaxID=145184;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=4998/89;  
 RA Greiser-Wilke I.M.;  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=4998/89;  
 RA Tajima M.;  
 RT "Prevalence of genotypes 1 and 2 of bovine viral diarrhea virus in  
 RT Lower Saxony, Germany.";  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ302959; CAC24757.1; -;  
 FT NON\_TER 1 1  
 FT NON\_TER 140 140  
 SQ SEQUENCE 140 AA; 15756 MW; 1A28D06B74391F74 CRC64;

Query Match 34.5%; Score 50; DB 12; Length 140;  
 Best Local Similarity 36.0%; Pred. No. 7.9;  
 Matches 9; Conservative 4; Mismatches 6; Indels 6; Gaps 1;  
 QY 1 CGEGYQSRVDHPHPFKPIVASITKC 25  
 Db 69 CGENFRKKEDLPHP-----IGKC 87  
 RESULT 14  
 Q95536 PRELIMINARY; PRT; 142 AA.  
 AC Q95536;  
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
 DE POPY-DRB5\*0603 protein (Fragment).  
 GN POPY-DRB5\*0603.  
 OS Pongo pygmaeus (Orangutan).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Pongo.  
 OX NCBI\_TaxID=9600;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93177102; PubMed=8439728;  
 RA Schonbach C., Vincek V., Mayer W.E., Golubic M., O'hugin C.,  
 RA Klein J.;  
 RT "Multiplication of Mhc-DRB5 loci in the orangutan: implications for  
 RT the evolution of DRB haplotypes.";  
 RL Mamm. Genome 4:159-170(1993).  
 DR EMBL; S56369; AAD13881.1; -;  
 DR HSSP; P13758; 1AOD.  
 DR InterPro; IPR007110; IG-like.  
 DR InterPro; IPR003597; IG-cl.  
 DR InterPro; IPR003006; IG\_MHC.  
 DR Pfam; PF00047; IG; 1.  
 DR SMART; SM00407; IGcl; 1.  
 DR PROSITE; PS0835; IG\_LIKE; 1.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 142 AA; 15525 MW; 02C2C3ABD44D063D CRC64;  
 Query Match 34.5%; Score 50; DB 7; Length 142;  
 Best Local Similarity 47.1%; Pred. No. 8;  
 Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;  
 QY 2 GEGYQSRVDHPHPFKPI 18  
 Db 73 GEGYTCQVEHPSPV 89  
 RESULT 15  
 Q9ESH2 PRELIMINARY; PRT; 314 AA.  
 AC Q9ESH2;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Preconditioning-inducible gene 1 protein.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Laser M., Li Y., Xu L., Darden A., Wu B.X., Hazard E.S. III,  
 RA Crosson C., Ma J.X.;  
 RT "Identification and characterization of a novel gene induced by  
 RT ischemic preconditioning in the retina.";  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF175224; AAG09182.1; -;  
 DR InterPro; IPR007135; Autophagy\_N.  
 DR InterPro; IPR007134; Autophagy\_N.

DR Pfam; PF03987; Autophagy\_C; 1.  
 DR Pfam; PF03986; Autophagy\_N; 1.  
 SQ SEQUENCE 314 AA; 35838 MW; 4C64E70F7E909BDF CRC64;  
 Query Match 34.5%; Score 50; DB 11; Length 314;  
 Best Local Similarity 43.8%; Pred. No. 18;  
 Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
 Qy 10 DHPHPKPIVRSITKC 25  
 Db 249 NHPHLPPLMCSVHPC 264

Search completed: February 26, 2004, 08:27:17  
 Job time : 34.4 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 08:16:50 ; Search time 46.6 Seconds  
(without alignments)  
151.581 Million cell updates/sec

Title: US-09-701-623c-7

Perfect score: 145

Sequence: 1 CGEGYQSRVDHPHPKPIVRSITKC 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq\_29Jan04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	145	100.0	25	3 AAY80000	Aay80000 Optimised
2	122	84.1	114	6 ABG74776	Abg74776 Rat Ige C
3	122	84.1	313	3 AAY79996	Aay79996 Rat Immun
4	122	84.1	340	3 AAB03643	Aab03643 Rat Ige h
5	122	84.1	341	3 AAB06205	Aab06205 Immunogen
6	122	84.1	346	6 ABG74785	Abg74785 Human CH2
7	122	84.1	428	6 ABP96589	Abp96589 Rat Ige h
8	121	83.4	45	3 AAY80018	Aay80018 Ige Immun
9	121	83.4	46	3 AAY80017	Aay80017 Ige Immun
10	106	73.1	25	3 AAY80001	Aay80001 Optimised
11	105	72.4	25	3 AAY79999	Aay79999 Optimised
12	105	72.4	45	3 AAY80019	Aay80019 Ige Immun
13	105	72.4	45	3 AAY80083	Aay80083 Ige Immun
14	105	72.4	45	3 AAY80020	Aay80020 Ige Immun
15	105	72.4	57	3 AAY80081	Aay80081 Ige Immun
16	105	72.4	62	3 AAY80080	Aay80080 Ige Immun
17	105	72.4	63	3 AAY80084	Aay80084 Ige Immun
18	103	71.0	25	3 AAY86602	Aay86602 Peptide s
19	103	71.0	25	3 AAY91212	Aay91212 Modified
20	103	71.0	25	3 AAY79998	Aay79998 Optimised
21	103	71.0	42	3 AAY86604	Aay86604 Peptide s
22	103	71.0	42	3 AAY91216	Aay91216 Modified
23	103	71.0	42	3 AAY91215	Aay91215 Modified
24	103	71.0	42	3 AAY91217	Aay91217 Modified
25	103	71.0	42	3 AAY80014	Aay80014 Ige Immun

26	103	71.0	44	7 ADD89950	Add89950 Ige pepti
27	103	71.0	45	3 AAY68605	Aay68605 Peptide s
28	103	71.0	45	3 AAY91218	Aay91218 Modified
29	103	71.0	45	3 AAY80007	Aay80007 Ige-CH3 d
30	103	71.0	45	7 ADD89951	Add89951 Ige pepti
31	103	71.0	46	3 AAY68603	Aay68603 Peptide s
32	103	71.0	46	3 AAY91213	Aay91213 Modified
33	103	71.0	46	3 AAY91214	Aay91214 Modified
34	103	71.0	46	3 AAY80011	Aay80011 Ige immun
35	103	71.0	56	3 AAY80016	Aay80016 Ige immun
36	103	71.0	59	3 AAY80010	Aay80010 Ige immun
37	103	71.0	60	3 AAY80013	Aay80013 Ige immun
38	103	71.0	60	3 AAY80015	Aay80015 Ige immun
39	103	71.0	63	3 AAY68606	Aay68606 Peptide s
40	103	71.0	63	3 AAY91219	Aay91219 Inv epito
41	103	71.0	63	3 AAY80008	Aay80008 Ige-CH3 d
42	103	71.0	63	3 AAY80012	Aay80012 Ige immun
43	98	67.6	332	5 AAU80298	Aau80298 Murine Ig
44	98	67.6	332	5 AAU80299	Aau80299 Murine Ig
45	98	67.6	332	5 AAU80297	Aau80297 Mouse Ige

## ALIGNMENTS

## RESULT 1

AAy80000  
ID AAY80000 standard; peptide; 25 AA.\*

XX AC AAY80000;

XX AC AAY80000;

DT 15-MAY-2000 (first entry)

XX DE Optimised Ige-CH3 domain antigen peptide for rat Ige.

XX Immunoglobulin E; Ige; epsilon heavy chain; antigenic; antigen;

KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;

KW antibody; allergy; allergic disease; immunisation; anti-allergy;

KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Rattus sp.

OS Synthetic.

XX WO9967293-A1.

XX 29-DEC-1999.

PD 21-JUN-1999; 99WO-US013959.

XX 20-JUN-1998; 98US-00100287.

PA (UNBI-) UNITED BIOMEDICAL INC.

PI Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions

PT for immunization against allergy.

XX Claim 1; Page 99; 155pp; English.

PS The present invention describes immunoglobulin E (Ige)-CH3 domain

CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic and

CC anti-asthmatic properties. (I) induces polyclonal antibodies specific for

CC a target effector site on the epsilon-heavy chain of Ige, and so

CC preventing triggering and activation of mast cells and basophils and

CC downregulation of Ige synthesis. Conjugates, or fusion peptides

CC containing (I) are used for active immunisation against Ige-mediated

CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy

CC dermatitis. Nucleic acids that encode these compounds are useful for

CC recombinant production of corresponding peptides or in DNA vaccines.

CC Conjugates of (I) that include a promiscuous T helper cell epitope

CC (functional in genetically diverse subjects), in addition to a B cell  
 CC target epitope, have increased immunogenicity and may include cyclic  
 CC constraints (disulfide bridge) to stabilise conformational features and  
 CC maximize cross-reactivity to the natural target. They induce safe (non-  
 CC anaphylactogenic) antibodies. AAY7994 to AAY8084 represent amino acid  
 CC sequences used in the exemplification of the present invention

XX Sequence 25 AA;

Query Match 100.0%; Score 145; DB 3; Length 25;  
 Best Local Similarity 100.0%; Pred. NO. 1.6e-14;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPIVRSITK 25

DB 1 CGEGYQSRVDHPHPKPIVRSITK 25

RESULT 2

ID ABG74776 standard; protein; 114 AA.

XX ABG74776;

XX 05-JUN-2003 (first entry)

XX Rat IgE CH3 domain.

XX CH3 domain; IgE; antigen; non-anaphylactic; anti-IgE; fusion protein;  
 XX dermatological; antiinflammatory; ophthalmological; allergy; asthma;  
 XX allergic rhinitis; gastrointestinal allergy; food allergy; eosinophilia;  
 XX conjunctivitis; glomerular nephritis; flea allergy; atopic dermatitis;  
 XX gene therapy; rat.

XX Rattus sp.

XX EPI262491-A2.

XX 04-DEC-2002.

XX 22-MAY-2002; 2002EP-00253606.

XX 22-MAY-2001; 2001US-0292638P.

XX (PFIZ ) PFIZER PROD INC.

XX Brown TM, Morsey MA;

XX WPI; 2003-122561/12.

XX N-PSDB; ACA55174.

XX Novel isolated antigenic peptide comprising amino acid residues of CH3  
 XX domain of IgE molecule from first species and a second unrelated species,  
 XX induces non-anaphylactic anti-IgE immune response in animal.

XX Claim 15; Page 31; 50pp; English.

XX This invention describes a novel antigenic peptide comprising amino acid  
 XX residues of an IgE CH3 domain from a first species (ADE1) and amino acid  
 XX residues of an IgE CH3 domain of a second unrelated species (ADE2), where  
 XX ADE1 is conserved in the IgE CH3 domain of the second species and ADE2 is  
 XX not conserved in the IgE CH3 domain of the first species. The novel  
 XX antigenic peptide induces a non-anaphylactic anti-IgE immune response in  
 XX an animal. The invention also discloses the polynucleotide sequence  
 XX encoding the antigenic peptide and an antigenic fusion protein comprising  
 XX the antigenic peptide of the invention and a heterologous protein  
 XX carrier, where the fusion protein induces an anti-IgE immune response  
 XX that does not cause anaphylaxis when administered to an animal. The  
 XX products of the invention have dermatological, antiinflammatory and  
 XX ophthalmological activity. The antigenic peptide described is capable of  
 XX preventing IgE from binding to high affinity receptors on mast cells and  
 XX basophils. The products of the invention are useful in the manufacture of  
 XX a medicament for treating or preventing IgE-mediated allergic disorders

CC including asthma, allergic rhinitis, gastrointestinal allergies such as  
 CC food allergies, eosinophilia, conjunctivitis, glomerular nephritis, flea  
 CC allergies or atopic dermatitis, in an animal, e.g. human or dog. The  
 CC polynucleotide products are useful for treating IgE-mediated allergic  
 CC disorders, by gene therapy. Antigenic peptides comprising conserved amino  
 CC acid residues of the CH3 domain of an IgE molecule from one species  
 CC flanked by variable amino acid residues of the CH3 domain of an IgE  
 CC molecule from a second unrelated species are capable of inducing a high  
 CC titre of anti-IgE antibodies when administered to an animal without  
 CC causing anaphylaxis. ABG74772-ABG74785 represents polypeptide sequences  
 CC used in designing the constructs described in the disclosure of the  
 CC invention

XX Sequence 114 AA;

Query Match 84.1%; Score 122; DB 6; Length 114;

Best Local Similarity 95.7%; Pred. NO. 2.4e-10;

Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPIVRSITK 24

DB 84 GEGYQSRVDHPHPKPIVRSITK 106

RESULT 3

AA79996

ID AAY79996 standard; protein; 313 AA.

XX AAY79996;

XX 15-MAY-2000 (first entry)

XX Rat immunoglobulin E epsilon heavy chain SEQ ID NO:3.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;  
 XX immunogenic; immunostimulatory; carrier protein; helper T cell epitope;  
 XX antibody; allergy; allergic disease; immunisation; anti-allergic;  
 XX anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Rattus sp.

XX WO9967293-A1.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US013959.

XX 20-JUN-1998; 98US-00100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions  
 XX for immunization against allergy.

XX Example 1; Page 66-68; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain  
 XX antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic and  
 XX anti-asthmatic properties. (I) induces polyclonal antibodies specific for  
 XX a target effector site on the epsilon-heavy chain of IgE, and so  
 XX preventing triggering and activation of mast cells and basophils and  
 XX downregulation of IgE synthesis. Conjugates, or fusion peptides,  
 XX containing (I) are used for active immunisation against IgE-mediated  
 XX allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy  
 XX dermatitis. Nucleic acids that encode these compounds are useful for  
 XX recombinant production of corresponding peptides or in DNA vaccines.  
 XX Conjugates of (I) that include a promiscuous T helper cell epitope  
 XX (functional in genetically diverse subjects), in addition to a B cell  
 XX target epitope, have increased immunogenicity and may include cyclic

CC constraints (disulfide bridge) to stabilise conformational features and  
 CC maximize cross-reactivity to the natural target. They induce safe (non-  
 CC anaphylactogenic) antibodies. AAY7994 to AAY8084 represent amino acid  
 CC sequences used in the exemplification of the present invention

XX SQ Sequence 313 AA;

Query Match 84.1%; Score 122; DB 3; Length 313;  
 Best Local Similarity 95.7%; Pred. No. 7.9e-10;  
 Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPIVRSITK 24  
 |||||  
 Db 190 GEGYQCRVDHPHPKPIVRSITK 212

RESULT 4

AAB03643  
 ID AAB03643 standard; protein; 340 AA.

XX AC AAB03643;

XX DT 22-NOV-2000 (first entry)

XX DE Rat IgE heavy chain constant regions 2, 3 and 4.

XX KW Rat; immunoglobulin E; IgE; vaccination; infection; allergy; asthma;  
 KW eczema; immunogenic peptide.

XX OS Rattus sp.

XX PN WO200025722-A2.

XX PD 11-MAY-2000.

XX PF 21-OCT-1999; 99WO-SE001896.

XX PR 02-NOV-1998; 98US-0106652P.

XX PR 22-SEP-1999; 99US-00401636.

XX PA (RESI-) RESISTENTIA PHARM AB.

XX PI Hellman LT;

XX PX WPI; 2000-365342/31.

XX PT Immunogenic polypeptides useful for preventing the harmful effects of  
 immunoglobulin E in mammals.

XX PS Disclosure; Fig 1; 50pp; English.

XX CC The present sequence is an immunogenic peptide consisting of the heavy  
 CC chain constant regions 2, 3 and 4 of the rat IgE. It was used to  
 CC construct a number of immunogenic peptides which consisted of regions of  
 CC IgE from different mammals, which appear to cause a stronger polyclonal  
 CC anti-self IgE response than peptides consisting of the same regions from  
 CC one mammal. Immunogenic peptides, particularly those consisting of  
 CC different heavy chain constant regions, can be used for vaccination in  
 CC humans, against bacterial and viral infections and allergies, such as  
 CC asthma, fur, pollen and food allergies and eczema

XX SQ Sequence 340 AA;

Query Match 84.1%; Score 122; DB 3; Length 340;  
 Best Local Similarity 95.7%; Pred. No. 7.8e-10;  
 Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPIVRSITK 24  
 |||||  
 Db 195 GEGYQCRVDHPHPKPIVRSITK 217

RESULT 5

AAB06206

ID AAB06206 standard; protein; 341 AA.

XX AC AAB06206;

XX DT 12-SEP-2003 (revised)

XX DT 22-NOV-2000 (first entry)

XX DE Immunogenic peptide consisting of opossum CH2, rat CH3 and opossum CH4.

XX KW Rat; opossum; immunoglobulin E; IgE; vaccination; infection; allergy;  
 KW asthma; eczema; immunogenic peptide.

XX OS Didelphis virginiana.

XX OS Rattus sp.

XX OS Chimeric.

XX PN WO200025722-A2.

XX PD 11-MAY-2000.

XX PF 21-OCT-1999; 99WO-SE001896.

XX PR 02-NOV-1998; 98US-0106652P.

XX PR 22-SEP-1999; 99US-00401636.

XX PA (RESI-) RESISTENTIA PHARM AB.

XX PI Hellman LT;

XX PX WPI; 2000-365342/31.

XX PT Immunogenic polypeptides useful for preventing the harmful effects of  
 immunoglobulin E in mammals.

XX PS Disclosure; Fig 2; 50pp; English.

XX CC The present sequence is an immunogenic peptide consisting of the heavy  
 CC chain constant regions 2 and 4 of the opossum IgE and the heavy chain  
 CC constant region 3 from the rat. It was shown to cause a stronger  
 CC polyclonal anti-self IgE response than peptides consisting of the same  
 CC regions from one mammal. Immunogenic peptides, particularly those  
 CC consisting of different heavy chain constant regions, can be used for  
 CC vaccination in humans, against bacterial and viral infections and  
 CC allergies, such as asthma, fur, pollen and food allergies and eczema.  
 CC (Updated on 12-SEP-2003 to standardise OS field)

XX SQ Sequence 341 AA;

Query Match 84.1%; Score 122; DB 3; Length 341;  
 Best Local Similarity 95.7%; Pred. No. 7.9e-10;  
 Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPIVRSITK 24  
 |||||  
 Db 206 GEGYQCRVDHPHPKPIVRSITK 228

RESULT 6  
 AEG74785

ID AEG74785 standard; protein; 346 AA.

XX AC AEG74785;

XX DT 05-JUN-2003 (first entry)

XX DE Human CH2-rat CH3-human CH4 IgE-5 fusion protein.

XX KW CH3 domain; IgE; antigen; non-anaphylactic; anti-IgE; fusion protein;  
 KW dermatological; antiinflammatory; ophthalmological; allergy; asthma;  
 KW allergic rhinitis; gastrointestinal allergy; food allergy; eosinophilia;  
 KW conjunctivitis; glomerular nephritis; flea allergy; atopic dermatitis;  
 KW gene therapy; human; rat.



AC AAY80018;  
 XX  
 DT 15-MAY-2000 (first entry)  
 XX  
 DE IGE immunogenic peptide conjugate SEQ ID NO:25.  
 XX  
 DE Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;  
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;  
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;  
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.  
 XX  
 OS Unidentified.  
 XX  
 XX WO9967293-A1.  
 XX  
 XX 29-DEC-1999.  
 XX  
 XX 21-JUN-1999; 99WO-US013959.  
 XX  
 XX 20-JUN-1998; 98US-00100287.  
 XX  
 XX (UNBI-) UNITED BIOMEDICAL INC.  
 XX  
 XX Wang CY, Walfield AM;  
 XX  
 XX WPI; 2000-160578/14.  
 XX  
 XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions  
 PT for immunization against allergy.  
 XX  
 XX Claim 14; Page 76; 155pp; English.  
 XX  
 XX The present invention describes immunoglobulin E (IgE)-CH3 domain  
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic and  
 CC anti-asthmatic properties. (I) induces polyclonal antibodies specific for  
 CC a target effector site on the epsilon-heavy chain of IgE, and so  
 CC preventing triggering and activation of mast cells and basophils and  
 CC downregulation of IgE synthesis. Conjugates, or fusion peptides,  
 CC containing (I) are used for active immunisation against IgE-mediated  
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy  
 CC dermatitis. Nucleic acids that encode these compounds are useful for  
 CC recombinant production of corresponding peptides or in DNA vaccines.  
 CC Conjugates of (I) that include a promiscuous T helper cell epitope  
 CC (functional in genetically diverse subjects), in addition to a B cell  
 CC target epitope, have increased immunogenicity and may include cyclic  
 CC constraints (disulfide bridge) to stabilise conformational features and  
 CC maximize cross-reactivity to the natural target. They induce safe (non-  
 CC anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino acid  
 CC sequences used in the exemplification of the present invention  
 XX  
 XX Sequence 45 AA;  
 SQ  
 Query Match 83.4%; Score 121; DB 3; Length 45;  
 Best Local Similarity 88.0%; Pred. No. 1.2e-10;  
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 CGEGYQSRVDHPHPKPIVRSITKC 25  
 DB 21 CGYGYQSIVDHPDPKPIVRSITKC 45  
 RESULT 9  
 AAY80017  
 ID AAY80017 standard; peptide; 46 AA.  
 XX  
 AC AAY80017;  
 XX  
 DT 15-MAY-2000 (first entry)  
 XX  
 DE IGE immunogenic peptide conjugate SEQ ID NO:24.  
 XX  
 DE Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;  
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;  
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;  
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

KW antibody; allergy; allergic disease; immunisation; anti-allergic;  
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.  
 OS Unidentified.  
 XX  
 XX WO9967293-A1.  
 XX  
 XX 29-DEC-1999.  
 XX  
 XX 21-JUN-1999; 99WO-US013959.  
 XX  
 XX 20-JUN-1998; 98US-00100287.  
 XX  
 XX (UNBI-) UNITED BIOMEDICAL INC.  
 XX  
 XX Wang CY, Walfield AM;  
 XX  
 XX WPI; 2000-160578/14.  
 XX  
 XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions  
 PT for immunization against allergy.  
 XX  
 XX Claim 14; Page 76; 155pp; English.  
 XX  
 XX The present invention describes immunoglobulin E (IgE)-CH3 domain  
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic and  
 CC anti-asthmatic properties. (I) induces polyclonal antibodies specific for  
 CC a target effector site on the epsilon-heavy chain of IgE, and so  
 CC preventing triggering and activation of mast cells and basophils and  
 CC downregulation of IgE synthesis. Conjugates, or fusion peptides,  
 CC containing (I) are used for active immunisation against IgE-mediated  
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy  
 CC dermatitis. Nucleic acids that encode these compounds are useful for  
 CC recombinant production of corresponding peptides or in DNA vaccines.  
 CC Conjugates of (I) that include a promiscuous T helper cell epitope  
 CC (functional in genetically diverse subjects), in addition to a B cell  
 CC target epitope, have increased immunogenicity and may include cyclic  
 CC constraints (disulfide bridge) to stabilise conformational features and  
 CC maximize cross-reactivity to the natural target. They induce safe (non-  
 CC anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino acid  
 CC sequences used in the exemplification of the present invention  
 XX  
 XX Sequence 46 AA;  
 SQ  
 Query Match 83.4%; Score 121; DB 3; Length 46;  
 Best Local Similarity 88.0%; Pred. No. 1.2e-10;  
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 CGEGYQSRVDHPHPKPIVRSITKC 25  
 DB 22 CGYGYQSIVDHPDPKPIVRSITKC 46  
 RESULT 10  
 AAY80001  
 ID AAY80001 standard; peptide; 25 AA.  
 XX  
 AC AAY80001;  
 XX  
 DT 15-MAY-2000 (first entry)  
 XX  
 DE Optimised IgE-CH3 domain antigen peptide for mouse IgE.  
 XX  
 DE Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;  
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;  
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;  
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.  
 XX  
 OS Mus sp.  
 OS Synthetic.  
 XX  
 XX WO9967293-A1.  
 XX



PD 29-DEC-1999.  
 XX 21-JUN-1999; 99WO-US013959.  
 XX 20-JUN-1998; 98US-00100287.  
 XX (UNBI-) UNITED BIOMEDICAL INC.  
 XX Wang CY, Walfield AM;  
 XX WPI; 2000-160578/14.  
 XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions  
 PT for immunization against allergy.  
 XX Claim 1; Page 100; 155pp; English.  
 XX The present invention describes immunoglobulin E (IgE)-CH3 domain  
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic and  
 CC anti-asthmatic properties. (I) induces polyclonal antibodies specific for  
 CC a target effector site on the epsilon-heavy chain of IgE, and so  
 CC preventing triggering and activation of mast cells and basophils and  
 CC downregulation of IgE synthesis. Conjugates, or fusion peptides,  
 CC containing (I) are used for active immunisation against IgE-mediated  
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy  
 CC dermatitis. Nucleic acids that encode these compounds are useful for  
 CC recombinant production of corresponding peptides or in DNA vaccines.  
 CC Conjugates of (I) that include a promiscuous T helper cell epitope  
 CC (functional in genetically diverse subjects), in addition to a B cell  
 CC target epitope, have increased immunogenicity and may include cyclic  
 CC constraints (disulfide bridge) to stabilise conformational features and  
 CC maximize cross-reactivity to the natural target. They induce safe (non-  
 CC anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino acid  
 CC sequences used in the exemplification of the present invention  
 XX Sequence 25 AA;  
 SQ  
 Query Match 73.1%; Score 106; DB 3; Length 25;  
 Best Local Similarity 80.0%; Pred. No. 1.1e-08;  
 Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 CGEGYQSRVDHPFPKPIVRSITKC 25  
 DB 1 CGYGYQSRVDHPFPKPIVRSITLC 25  
 RESULT 11  
 ID AAY79999 standard; peptide; 25 AA.  
 AC AAY79999;  
 XX 15-MAY-2000 (first entry)  
 DE Optimised IgE-CH3 domain antigen peptide for dog IgE.  
 XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;  
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;  
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;  
 XX anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.  
 OS Canis sp.  
 OS Synthetic.  
 XX WO9967293-A1.  
 XX 29-DEC-1999.  
 XX 21-JUN-1999; 99WO-US013959.  
 XX 20-JUN-1998; 98US-00100287.  
 XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfield AM;  
 XX WPI; 2000-160578/14.  
 XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions  
 PT for immunization against allergy.  
 XX Claim 1; Page 99; 155pp; English.  
 XX The present invention describes immunoglobulin E (IgE)-CH3 domain  
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic and  
 CC anti-asthmatic properties. (I) induces polyclonal antibodies specific for  
 CC a target effector site on the epsilon-heavy chain of IgE, and so  
 CC preventing triggering and activation of mast cells and basophils and  
 CC downregulation of IgE synthesis. Conjugates, or fusion peptides,  
 CC containing (I) are used for active immunisation against IgE-mediated  
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy  
 CC dermatitis. Nucleic acids that encode these compounds are useful for  
 CC recombinant production of corresponding peptides or in DNA vaccines.  
 CC Conjugates of (I) that include a promiscuous T helper cell epitope  
 CC (functional in genetically diverse subjects), in addition to a B cell  
 CC target epitope, have increased immunogenicity and may include cyclic  
 CC constraints (disulfide bridge) to stabilise conformational features and  
 CC maximize cross-reactivity to the natural target. They induce safe (non-  
 CC anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino acid  
 CC sequences used in the exemplification of the present invention  
 XX Sequence 25 AA;  
 SQ  
 Query Match 72.4%; Score 105; DB 3; Length 25;  
 Best Local Similarity 76.0%; Pred. No. 1.6e-08;  
 Matches 19; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 CGEGYQSRVDHPFPKPIVRSITKC 25  
 DB 1 CGETYSRVTHPLPKDIVRSIAKC 25  
 RESULT 12  
 ID AAY80019 standard; peptide; 45 AA.  
 AC AAY80019;  
 XX 15-MAY-2000 (first entry)  
 DE IgE immunogenic peptide conjugate SEQ ID NO:26.  
 XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;  
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;  
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;  
 XX anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.  
 OS Unidentified.  
 OS WO9967293-A1.  
 XX 29-DEC-1999.  
 XX 21-JUN-1999; 99WO-US013959.  
 XX 20-JUN-1998; 98US-00100287.  
 XX (UNBI-) UNITED BIOMEDICAL INC.  
 XX Wang CY, Walfield AM;  
 XX WPI; 2000-160578/14.  
 XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions  
 PT for immunization against allergy.

AA CC The present invention describes immunoglobulin E (IgE)-CH3 domain  
CC CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic and  
CC CC anti-asthmatic properties. (I) induces polyclonal antibodies specific for  
CC CC a target effector site on the epsilon-heavy chain of IgE, and so  
CC CC preventing triggering and activation of mast cells and basophils and  
CC CC downregulation of IgE synthesis. Conjugates, or fusion peptides,

The present invention describes immunoglobulin E (IgE)-CH3 domain antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic and anti-asthmatic properties. (I) induces polyclonal antibodies specific for a target effector site on the epsilon-heavy chain of IgE, and so preventing triggering and activation of mast cells and basophils and downregulation of IgE synthesis. Conjugates, or fusion peptides, containing (I) are used for active immunisation against IgE-mediated allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy dermatitis. Nucleic acids that encode these compounds are useful for recombinant production of corresponding peptides or in DNA vaccines. Conjugates of (I) that include a promiscuous T helper cell epitope (functional in genetically diverse subjects), in addition to a B cell target epitope, have increased immunogenicity and may include cyclic constraints (disulfide bridge) to stabilise conformational features and

Query Match 72.4%; Score 105; DB 3;  
Best Local Similarity 76.0%; Pred. No. 3.9e-08;

```
; ORGANISM: Felis catus
US-10-409-772-2
Query Match      68.1%; Score 92; DB 15; Length 496;
Best Local Similarity 69.6%; Pred. No. 1.8e-05;
Matches 16; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY      2 GETYKSTVSHDPLPREVVRSAK 24
      |||||:|||||:|||||:|||||
Db      359 GETYQKVTHTPLPKDIVRSAK 381

RESULT 12
US-10-409-772-29
; Sequence 29, Application US/10409772
; Publication No. US2003021856A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine
; APPLICANT: Weber, Eric
; TITLE OF INVENTION: Feline Immunoglobulin E Molecules and Related Methods
; FILE REFERENCE: P-1047
; CURRENT APPLICATION NUMBER: US/10/409,772
; CURRENT FILING DATE: 2003-04-07
; PRIOR APPLICATION NUMBER: US/09/479,614
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Felis catus
US-10-409-772-29
Query Match      68.1%; Score 92; DB 15; Length 496;
Best Local Similarity 69.6%; Pred. No. 1.8e-05;
Matches 16; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY      2 GETYKSTVSHDPLPREVVRSAK 24
      |||||:|||||:|||||:|||||
Db      359 GETYQKVTHTPLPKDIVRSAK 381

RESULT 13
US-10-152-190-1
; Sequence 1, Application US/10152190
; Publication No. US20030096369A1
; GENERAL INFORMATION:
; APPLICANT: Morsey, Mohamad A.
; TITLE OF INVENTION: No. US20030096369A1-anaphylactogenic Ige vaccines
; FILE REFERENCE: PC11011A
; CURRENT APPLICATION NUMBER: US/10/152,190
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Dog CH3 domain
US-10-152-190-1
Query Match      60.7%; Score 82; DB 14; Length 114;
Best Local Similarity 65.2%; Pred. No. 0.00011;
Matches 15; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY      2 GETYKSTVSHDPLPREVVRSAK 24
      |||||:|||||:|||||:|||||
Db      84 GETYICRVTHPLPKDIVRSAK 106

RESULT 14
US-10-152-190-3
; Sequence 3, Application US/10152190
; Publication No. US20030096369A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Morsey, Mohamad A.
; TITLE OF INVENTION: No. US20030096369A1-anaphylactogenic Ige vaccines
; FILE REFERENCE: PC11011A
; CURRENT APPLICATION NUMBER: US/10/152,190
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Human CH3/dog CH3 domain chimera
US-10-152-190-3
Query Match      60.7%; Score 82; DB 14; Length 115;
Best Local Similarity 65.2%; Pred. No. 0.00011;
Matches 15; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY      2 GETYKSTVSHDPLPREVVRSAK 24
      |||||:|||||:|||||:|||||
Db      85 GETYICRVTHPLPKDIVRSAK 107

RESULT 15
US-10-152-190-2
; Sequence 2, Application US/10152190
; Publication No. US20030096369A1
; GENERAL INFORMATION:
; APPLICANT: Morsey, Mohamad A.
; TITLE OF INVENTION: No. US20030096369A1-anaphylactogenic Ige vaccines
; FILE REFERENCE: PC11011A
; CURRENT APPLICATION NUMBER: US/10/152,190
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Human CH3/dog CH3 domain fusion
US-10-152-190-2
Query Match      60.7%; Score 82; DB 14; Length 117;
Best Local Similarity 65.2%; Pred. No. 0.00011;
Matches 15; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY      2 GETYKSTVSHDPLPREVVRSAK 24
      |||||:|||||:|||||:|||||
Db      87 GETYICRVTHPLPKDIVRSAK 109

Search completed: February 26, 2004, 08:46:54
Job time : 26.2 secs
```

```
QY      2 GETYKSTVSHDPDLPREVVRSAK 24
      ||||: |:||||:|:||||:|
Db      294 GETYQCKVTHPDLPKDIVRSAK 316

RESULT 7
US-10-409-772-14
; Sequence 14, Application US/10409772
; Publication No. US20030216565A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine
; APPLICANT: Weber, Eric
; TITLE OF INVENTION: Feline Immunoglobulin E Molecules and Related Methods
; FILE REFERENCE: P-1047
; CURRENT APPLICATION NUMBER: US/10/409,772
; PRIOR FILING DATE: 2003-04-07
; PRIOR APPLICATION NUMBER: US/09/479,614
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Felis catus
US-10-409-772-14
Query Match      68.1%; Score 92; DB 15; Length 431;
Best Local Similarity 69.6%; Pred. No. 1.5e-05;
Matches 16; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY      2 GETYKSTVSHDPDLPREVVRSAK 24
      ||||: |:||||:|:||||:|
Db      294 GETYQCKVTHPDLPKDIVRSAK 316

RESULT 8
US-09-479-614-2
; Sequence 2, Application US/09479614
; Publication No. US2003001183A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine
; APPLICANT: Weber, Eric
; TITLE OF INVENTION: Feline Immunoglobulin E Molecules and Related Methods
; FILE REFERENCE: P-1047
; CURRENT APPLICATION NUMBER: US/09/479,614
; CURRENT FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 60/115,033
; EARLIER FILING DATE: 1999-01-07
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Felis catus
US-09-479-614-2
Query Match      68.1%; Score 92; DB 10; Length 496;
Best Local Similarity 69.6%; Pred. No. 1.8e-05;
Matches 16; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY      2 GETYKSTVSHDPDLPREVVRSAK 24
      ||||: |:||||:|:||||:|
Db      359 GETYQCKVTHPDLPKDIVRSAK 381

RESULT 9
US-09-479-614-29
; Sequence 29, Application US/09479614
; Publication No. US20030013183A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine
; APPLICANT: Weber, Eric
```

```
; TITLE OF INVENTION: Feline Immunoglobulin E Molecules and Related Methods
; FILE REFERENCE: P-1047
; CURRENT APPLICATION NUMBER: US/09/479,614
; CURRENT FILING DATE: 2000-01-07
; EARLIER APPLICATION NUMBER: 60/115,033
; EARLIER FILING DATE: 1999-01-07
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Felis catus
US-09-479-614-29
Query Match      68.1%; Score 92; DB 10; Length 496;
Best Local Similarity 69.6%; Pred. No. 1.8e-05;
Matches 16; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY      2 GETYKSTVSHDPDLPREVVRSAK 24
      ||||: |:||||:|:||||:|
Db      359 GETYQCKVTHPDLPKDIVRSAK 381

RESULT 10
US-10-214-524-25
; Sequence 25, Application US/10214524
; Publication No. US20030073142A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Swei-Shen Alex
; APPLICANT: Yang, Yong-Min
; APPLICANT: Barankiewicz, Theresa J.
; APPLICANT: Chen, Zhong
; TITLE OF INVENTION: IMMUNOGLOBULIN E VACCINES AND METHODS OF USE THEREOF
; FILE REFERENCE: IGE-00101.P.1.1
; CURRENT APPLICATION NUMBER: US/10/214,524
; CURRENT FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/312,120
; PRIOR FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Cat (Felis catus)
US-10-214-524-25
Query Match      68.1%; Score 92; DB 14; Length 496;
Best Local Similarity 69.6%; Pred. No. 1.8e-05;
Matches 16; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY      2 GETYKSTVSHDPDLPREVVRSAK 24
      ||||: |:||||:|:||||:|
Db      359 GETYQCKVTHPDLPKDIVRSAK 381

RESULT 11
US-10-409-772-2
; Sequence 2, Application US/10409772
; Publication No. US20030216565A1
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine
; APPLICANT: Weber, Eric
; TITLE OF INVENTION: Feline Immunoglobulin E Molecules and Related Methods
; FILE REFERENCE: P-1047
; CURRENT APPLICATION NUMBER: US/10/409,772
; CURRENT FILING DATE: 2003-04-07
; PRIOR APPLICATION NUMBER: US/09/479,614
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 496
; TYPE: PRT
```

;; TITLE OF INVENTION: Stabilized Synthetic Immunogen Delivery System  
;; FILE REFERENCE: Immunogen Delivery System  
;; CURRENT APPLICATION NUMBER: US/10/076,674  
;; CURRENT FILING DATE: 2002-04-23  
;; NUMBER OF SEQ ID NOS: 11  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 10  
;; LENGTH: 44  
;; TYPE: PRT  
;; ORGANISM: Human  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (19)..(19)  
;; OTHER INFORMATION: Xaa indicates epsilon-Lys  
US-10-076-674-10

Query Match 69.6%; Score 94; DB 14; Length 44;  
Best Local Similarity 64.0%; Pred. No. 5.4e-07;  
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGETYKSTVSHPDLPREVVRSIAC 25  
Db 20 CGETYQSRVTHPHLPALMRSTTKC 44

## RESULT 3

US-10-355-161A-10  
;; Sequence 10, Application US/10355161A  
;; Publication No. US20040009897A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Sokoll, Kenneth K.

;; TITLE OF INVENTION: Stabilized Synthetic Immunogen Delivery System  
;; FILE REFERENCE: Immunogen Delivery System  
;; CURRENT APPLICATION NUMBER: US/10/355,161A  
;; CURRENT FILING DATE: 2003-01-31  
;; PRIOR APPLICATION NUMBER: US 10/076674  
;; PRIOR FILING DATE: 2002-02-14  
;; NUMBER OF SEQ ID NOS: 13  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 10  
;; LENGTH: 44  
;; TYPE: PRT  
;; ORGANISM: Human  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (19)..(19)  
;; OTHER INFORMATION: Xaa indicates epsilon-Lys  
US-10-355-161A-10

Query Match 69.6%; Score 94; DB 15; Length 44;  
Best Local Similarity 64.0%; Pred. No. 5.4e-07;  
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGETYKSTVSHPDLPREVVRSIAC 25  
Db 20 CGETYQSRVTHPHLPALMRSTTKC 44

## RESULT 4

US-10-076-674-11  
;; Sequence 11, Application US/10076674  
;; Publication No. US20030165478A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Sokoll, Kenneth K.

;; TITLE OF INVENTION: Stabilized Synthetic Immunogen Delivery System  
;; FILE REFERENCE: Immunogen Delivery System  
;; CURRENT APPLICATION NUMBER: US/10/076,674  
;; CURRENT FILING DATE: 2002-04-23  
;; NUMBER OF SEQ ID NOS: 11  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 11  
;; LENGTH: 45  
;; TYPE: PRT

;; ORGANISM: Human  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (20)..(20)  
;; OTHER INFORMATION: Xaa indicates epsilon-Lys  
US-10-076-674-11

Query Match 69.6%; Score 94; DB 14; Length 45;  
Best Local Similarity 64.0%; Pred. No. 5.5e-07;  
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGETYKSTVSHPDLPREVVRSIAC 25  
Db 21 CGETYQSRVTHPHLPALMRSTTKC 45

## RESULT 5

US-10-355-161A-11  
;; Sequence 11, Application US/10355161A  
;; Publication No. US20040009897A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Sokoll, Kenneth K.  
;; TITLE OF INVENTION: Stabilized Synthetic Immunogen Delivery System  
;; FILE REFERENCE: Immunogen Delivery System  
;; CURRENT APPLICATION NUMBER: US/10/355,161A  
;; CURRENT FILING DATE: 2003-01-31  
;; PRIOR APPLICATION NUMBER: US 10/076674  
;; PRIOR FILING DATE: 2002-02-14  
;; NUMBER OF SEQ ID NOS: 13  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 11  
;; LENGTH: 45  
;; TYPE: PRT  
;; ORGANISM: Human  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (20)..(20)  
;; OTHER INFORMATION: Xaa indicates epsilon-Lys  
US-10-355-161A-11

Query Match 69.6%; Score 94; DB 15; Length 45;  
Best Local Similarity 64.0%; Pred. No. 5.5e-07;  
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGETYKSTVSHPDLPREVVRSIAC 25  
Db 21 CGETYQSRVTHPHLPALMRSTTKC 45

## RESULT 6

US-09-479-614-14  
;; Sequence 14, Application US/09479614  
;; Publication No. US20030013183A1  
;; GENERAL INFORMATION:  
;; APPLICANT: McCall, Catharine  
;; APPLICANT: Weber, Eric  
;; TITLE OF INVENTION: Peline Immunoglobulin E Molecules and Related Methods  
;; FILE REFERENCE: P-1047  
;; CURRENT APPLICATION NUMBER: US/09/479,614  
;; CURRENT FILING DATE: 2000-01-07  
;; EARLIER APPLICATION NUMBER: 60/115,033  
;; EARLIER FILING DATE: 1999-01-07  
;; NUMBER OF SEQ ID NOS: 34  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 14  
;; LENGTH: 431  
;; TYPE: PRT  
;; ORGANISM: Felis catus  
US-09-479-614-14

Query Match 68.1%; Score 92; DB 10; Length 431;  
Best Local Similarity 69.6%; Pred. No. 1.5e-05;  
Matches 16; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 08:27:32 ; Search time 26.2 Seconds

(without alignments)  
201.482 Million cell updates/sec

Title: US-09-701-623C-84

Perfect score: 135

Sequence: 1 CGTYKSTVSHDPLPREVRSIAKC 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2\_6/ptodata/2/pubpaa/PCT NEW PUB. pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT NEW PUB. pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06 NEW PUB. pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06 PUBCOMB. pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US06 NEW PUB. pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US06 PUBCOMB. pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08 NEW PUB. pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08 PUBCOMB. pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A PUBCOMB. pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B PUBCOMB. pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C PUBCOMB. pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09 NEW PUB. pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A PUBCOMB. pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B PUBCOMB. pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C PUBCOMB. pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10 NEW PUB. pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60 NEW PUB. pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60 PUBCOMB. pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	112	83.0	569	14	US-10-214-524-30
2	94	69.6	44	14	US-10-076-674-10
3	94	69.6	44	15	US-10-355-161A-10
4	94	69.6	45	14	US-10-076-674-11
5	94	69.6	45	15	US-10-355-161A-11
6	92	68.1	431	10	US-09-479-614-14
7	92	68.1	431	15	US-10-409-772-14
8	92	68.1	496	10	US-09-479-614-29
9	92	68.1	496	15	US-10-409-772-29
10	92	68.1	496	14	US-10-214-524-25
11	92	68.1	496	15	US-10-409-772-2
12	92	68.1	496	15	US-10-409-772-29
13	82	60.7	114	14	US-10-152-190-1
14	82	60.7	115	14	US-10-152-190-3
15	82	60.7	117	14	US-10-152-190-2

16	82	60.7	341	9	US-09-401-636-11	Sequence 11, Appl
17	82	60.7	341	14	US-10-176-664-11	Sequence 11, Appl
18	82	60.7	346	14	US-10-152-190-10	Sequence 10, Appl
19	82	60.7	347	14	US-10-152-190-12	Sequence 12, Appl
20	82	60.7	348	14	US-10-152-190-11	Sequence 11, Appl
21	82	60.7	426	14	US-10-214-524-28	Sequence 28, Appl
22	81	60.0	345	9	US-09-401-636-10	Sequence 10, Appl
23	81	60.0	345	14	US-10-176-664-10	Sequence 10, Appl
24	81	60.0	567	14	US-10-214-524-33	Sequence 33, Appl
25	74	54.8	426	14	US-10-214-524-26	Sequence 26, Appl
26	74	54.8	563	14	US-10-214-524-35	Sequence 35, Appl
27	71	52.6	115	14	US-10-152-190-4	Sequence 4, Appl
28	71	52.6	129	14	US-10-152-190-6	Sequence 6, Appl
29	71	52.6	222	9	US-09-809-746-2	Sequence 2, Appl
30	71	52.6	222	10	US-09-809-715-6	Sequence 6, Appl
31	71	52.6	320	10	US-09-847-208-6	Sequence 6, Appl
32	71	52.6	323	9	US-09-949-375A-4	Sequence 2, Appl
33	71	52.6	323	9	US-09-949-375A-4	Sequence 4, Appl
34	71	52.6	323	9	US-09-949-375A-6	Sequence 6, Appl
35	71	52.6	330	9	US-09-949-375A-10	Sequence 10, Appl
36	71	52.6	331	9	US-09-401-636-1	Sequence 1, Appl
37	71	52.6	331	14	US-10-176-664-1	Sequence 1, Appl
38	71	52.6	331	14	US-10-207-655-329	Sequence 329, App
39	71	52.6	336	9	US-09-949-375A-8	Sequence 8, Appl
40	71	52.6	342	9	US-09-401-636-8	Sequence 8, Appl
41	71	52.6	342	14	US-10-176-664-8	Sequence 8, Appl
42	71	52.6	347	14	US-10-152-190-13	Sequence 13, Appl
43	71	52.6	426	14	US-10-214-524-27	Sequence 27, Appl
44	71	52.6	427	10	US-09-847-208-5	Sequence 5, Appl
45	71	52.6	428	9	US-09-916-230-1	Sequence 1, Appl

# ALIGNMENTS

## RESULT 1

US-10-214-524-30  
; Sequence 30, Application US/10214524  
; Publication No. US20030073142A1  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Swei-Shen Alex  
; APPLICANT: Yang, Yong-Min  
; APPLICANT: Barankiewicz, Theresa J.  
; APPLICANT: Chen, Zhong  
; TITLE OF INVENTION: IMMUNOGLOBULIN E VACCINES AND METHODS OF USE THEREOF  
; FILE REFERENCE: IGE-00101.P.1.1  
; CURRENT APPLICATION NUMBER: US/10/214,524  
; CURRENT FILING DATE: 2002-08-08  
; PRIOR APPLICATION NUMBER: 60/312,120  
; PRIOR FILING DATE: 2001-08-13  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 30  
; LENGTH: 569  
; TYPE: PRT  
; ORGANISM: Horse (Equus caballus)  
US-10-214-524-30

Query Match 83.0%; Score 112; DB 14; Length 569;  
Best Local Similarity 95.7%; Pred. No. 1.9e-08;  
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GETYKSTVSHDPLPREVRSIAK 24

DB 433 GETYKSTVSHDPLPREVRSIAK 455

## RESULT 2

US-10-076-674-10  
; Sequence 10, Application US/10076674  
; Publication No. US20030165478A1  
; GENERAL INFORMATION:  
; APPLICANT: Sokoll, Kenneth K.

```

XX PD 29-DEC-1999.
XX FT /label= Gly, Thr
XX FT Misc-difference 10
XX FT /label= His, Thr
XX FT Misc-difference 11
XX FT /label= Lys, Arg
XX FT Misc-difference 12
XX FT /label= Ile, Met, Leu
XX FT Misc-difference 14
XX FT /label= Gly, Thr
XX FT Misc-difference 15
XX FT /label= Ile, Met, Val
XX XX
XX PN WO9966952-A1.
XX XX
XX PD 29-DEC-1999.
XX XX
XX PF 21-JUN-1999; 99WO-US013960.
XX XX
XX PR 20-JUN-1998; 98US-00100414.
XX PA (UNBI-) UNITED BIOMEDICAL INC.
XX PI Wang CY;
XX XX
XX DR WPI; 2000-160562/14.
XX XX
XX PT New peptide immunogen containing luteinizing hormone-releasing hormone
XX PT antigen site and helper T cell epitope, for e.g. contraception and
XX PT treatment of cancer.
XX PS Disclosure; Page 95; 102pp; English.
XX CC The specification describes peptide immunogens comprising a synthetic
XX CC helper T cell (TH) epitope and a target antigen, luteinising hormone-
XX CC releasing hormone (LHRH). The peptide immunogens cause induction of a
XX CC specific immune response to LHRH which is involved in regulation of
XX CC spermatogenesis, ovulation, oestrus, sexual development and secretion of
XX CC sex hormones. Provision of a promiscuous T helper epitope (which is
XX CC functional in genetically diverse subjects) provides optimum
XX CC immunogenicity to the B cell epitopes of the target antigen and thus high
XX CC antibody titres against the target antigen. The peptide immunogens of the
XX CC invention are used to vaccinate against mammalian LHRH, for use as
XX CC (reversible) contraceptive; control of hormone-dependent tumours (cancer
XX CC of prostate or breast, also endometriosis); to prevent boar taint (and
XX CC improve meat quality) and for immunocastration. The present sequence
XX CC appears in the specification
XX SQ
XX Query Match 69.6%; Score 94; DB 3; Length 25;
XX Best Local Similarity 64.0%; Pred. No. 2.5e-07;
XX Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
QY 1 CGETYKSTVSHPDLPREVVRSIAC 25
DB 1 CGETYQSRVTHPHLPALMRSTTKC 25
XX
RESULT 15
AAY68604
XX ID AAY68604 standard; peptide; 42 AA.
XX AC AAY68604;
XX XX
XX DT 05-MAY-2000 (first entry)
XX DE Peptide sequence of the invention.
XX XX
XX KW Helper T cell epitope; peptide immunogen; LHRH;
XX KW luteinising hormone-releasing hormone; spermatogenesis; ovulation;
XX KW oestrus; sexual development; sex hormone; promiscuous T helper epitope;
XX KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
XX KW breast cancer; endometriosis; boar taint; meat quality; immunocastration.
XX OS Unidentified.
XX XX
XX FH Key Location/Qualifiers
XX FT Misc-difference 1 /label= Ile, Met, Leu
XX FT Misc-difference 2 /label= Ser, Thr
XX FT Misc-difference 5 /label= Lys, Arg
XX FT Misc-difference 6

```

```

FT FT /label= Gly, Thr
FT FT Misc-difference 10
FT FT /label= His, Thr
FT FT Misc-difference 11
FT FT /label= Lys, Arg
FT FT Misc-difference 12
FT FT /label= Ile, Met, Leu
FT FT Misc-difference 14
FT FT /label= Gly, Thr
FT FT Misc-difference 15
FT FT /label= Ile, Met, Val
XX XX
XX PN WO9966952-A1.
XX XX
XX PD 29-DEC-1999.
XX XX
XX PF 21-JUN-1999; 99WO-US013960.
XX XX
XX PR 20-JUN-1998; 98US-00100414.
XX PA (UNBI-) UNITED BIOMEDICAL INC.
XX PI Wang CY;
XX XX
XX DR WPI; 2000-160562/14.
XX XX
XX PT New peptide immunogen containing luteinizing hormone-releasing hormone
XX PT antigen site and helper T cell epitope, for e.g. contraception and
XX PT treatment of cancer.
XX PS Disclosure; Page 95; 102pp; English.
XX CC The specification describes peptide immunogens comprising a synthetic
XX CC helper T cell (TH) epitope and a target antigen, luteinising hormone-
XX CC releasing hormone (LHRH). The peptide immunogens cause induction of a
XX CC specific immune response to LHRH which is involved in regulation of
XX CC spermatogenesis, ovulation, oestrus, sexual development and secretion of
XX CC sex hormones. Provision of a promiscuous T helper epitope (which is
XX CC functional in genetically diverse subjects) provides optimum
XX CC immunogenicity to the B cell epitopes of the target antigen and thus high
XX CC antibody titres against the target antigen. The peptide immunogens of the
XX CC invention are used to vaccinate against mammalian LHRH, for use as
XX CC (reversible) contraceptive; control of hormone-dependent tumours (cancer
XX CC of prostate or breast, also endometriosis); to prevent boar taint (and
XX CC improve meat quality) and for immunocastration. The present sequence
XX CC appears in the specification
XX SQ
XX Query Match 69.6%; Score 94; DB 3; Length 42;
XX Best Local Similarity 64.0%; Pred. No. 2.5e-07;
XX Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
QY 1 CGETYKSTVSHPDLPREVVRSIAC 25
DB 18 CGETYQSRVTHPHLPALMRSTTKC 42
XX
Search completed: February 26, 2004, 08:23:35
Job time : 46.6 secs

```



CC helper T cell (Th) epitope and a target antigen, luteinising hormone-releasing hormone (LHRH). The peptide immunogens cause induction of a specific immune response to LHRH which is involved in regulation of spermatogenesis, ovulation of a primate, sexual development and secretion of sex hormones. Provision of a promiscuous T helper epitope (which is functional in genetically diverse subjects) provides optimum immunogenicity to the B cell epitopes of the target antigen and thus high antibody titres against the target antigen. The peptide immunogens of the invention are used to vaccinate against mammalian LHRH, for use as (reversible) contraceptive; control of hormone-dependent tumours (cancer of prostate or breast, also endometriosis); to prevent boar taint (and improve meat quality) and for immunocastration. The present sequence appears in the specification

XX Sequence 25 AA;

Query Match 69.6%; Score 94; DB 3; Length 25;  
Best Local Similarity 64.0%; Pred. No. 1.3e-07;  
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGETYKSTVSHPDLPREVVRSIAC 25

DB 1 CGETYQSRVTHPLPALMRSTTKC 25

RESULT 13

AAAY91212  
ID AA91212 standard; peptide; 25 AA.

XX AC AA91212;

DT 22-MAY-2000 (first entry)

DE Modified human IgE CH3 domain, SEQ ID NO:92.

XX Promiscuous T-cell epitope; measles virus F protein; MVF;  
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;  
KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;  
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;  
KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;  
KW Plasmodium falciparum; circumsporozoite; antimalarial; CERP;  
KW cholesteryl ester transport protein; anti-arteriosclerotic.

XX Homo sapiens.

OS Synthetic.

XX WO966957-A2.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US013975.

XX 20-JUN-1998; 98US-00100412.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang.CY;

XX WPI; 2000-160564/14.

XX New artificial T helper cell epitope and derived immunogens with target antigenic site, for immunization against e.g. malaria, arteriosclerosis or human immune deficiency virus.

XX Example 6; Page 40; 129pp; English.

XX The invention relates to novel promiscuous T helper cell epitopes (Th),  
CC and immunogenic peptides comprising the Th epitopes of the invention  
CC along with B cell epitopes. The Th epitopes and peptide immunogens  
CC containing them, are used to induce a T helper cell response,  
CC specifically against Plasmodium falciparum, cholesteryl ester transport  
CC protein (CERP) or HIV epitopes, but more generally against any pathogen,  
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and

CC peptide immunogens may be used for prevention and/or treatment of  
CC infections (HIV, foot-and-mouth disease or malaria); for cancer  
CC immunotherapy; for inhibition of the action of luteinising hormone  
CC releasing hormone (LHRH) for contraception, treatment of hormone-  
CC dependent cancer, prevention of boar taint in meat, and immunocastration)  
CC ; for promoting the growth of animals; or for treating allergies or  
CC arteriosclerosis. Incorporation of a promiscuous Th (functional in  
CC genetically diverse subjects) into an immunogen improves capacity to  
CC induce a strong T helper cell-mediated immune response, resulting in  
CC production of antibodies against a target antigen. Th can replace carrier  
CC proteins and pathogen-derived T helper epitopes. Sequence AA91121  
CC represents a promiscuous T helper epitope from the measles virus F (MVF)  
CC protein and sequences AA91122-Y91142, AA91226 and AA91245-Y91246  
CC represent synthetic Th epitopes based on the MVF Th epitope. Sequence  
CC AA91143 represents a promiscuous Th epitope from hepatitis B virus (HBV)  
CC surface antigen, and sequences AA91144-Y91155 are synthetic epitopes  
CC derived from this HBV epitope. AA91156-Y91196, AA91227 and AA91242-  
CC Y91244 are antigenic peptides comprising an LHRH sequence joined to a  
CC promiscuous Th epitope. AA91197 is the LHRH target antigenic peptide  
CC used in these LHRH antigenic peptides. AA91200 is somatostatin, and  
CC AA91201-Y91207 are antigenic peptides comprising somatostatin and a Th  
CC epitope. Somatostatin immunogens may be used to promote growth in  
CC livestock. AA91208 is a human CD4 CDR2-like domain antigenic site, and  
CC AA91209-Y90211 are MVF Th epitope/CD4 CDR2 antigenic peptides which may  
CC be used to prevent HIV infection of T cells. AA90212 is a modified  
CC version of a human IgE (immunoglobulin E) CH3 domain, and AA90213-Y90219  
CC are Th epitope/IgE CH3 antigenic peptides which may be used in the  
CC treatment of allergies. AA91220 is a peptide derived from foot and mouth  
CC disease virus (FMDV) VP1 capsid protein and AA91221-Y91222 comprise this  
CC peptide and a Th epitope. AA91223 is a Plasmodium falciparum  
CC circumsporozoite (CS) target antigen, and AA91224-Y91225 comprise the CS  
CC antigen and an MVF Th epitope and may be used in a malaria vaccine.  
CC AA91228-Y91231 represent CERP-derived peptides and AA91232-Y91241 are  
CC immunogens comprising a CERP peptide and a Th epitope which may be used  
CC to prevent or treat arteriosclerosis and cardiovascular disease. AA91247  
CC and AA91252-Y91257 are HIV-1 neutralising B-cell epitopes, and AA91248-  
CC Y91251 and AA91258-Y91273 are antigenic peptides comprising MVF Th and  
CC HIV-1 B-cell epitope which may be used as a component in an anti-HIV-1  
CC vaccine. AA91198 and AA91199 are respectively an immunostimulatory  
CC inactivating protein epitope from *Yersinia* species, and hinge spacer peptide,  
CC both of which may optionally be used in the antigenic peptides of the  
CC invention

XX SQ Sequence 25 AA;

Query Match 69.6%; Score 94; DB 3; Length 25;

Best Local Similarity 64.0%; Pred. No. 1.3e-07;

Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGETYKSTVSHPDLPREVVRSIAC 25

DB 1 CGETYQSRVTHPLPALMRSTTKC 25

RESULT 14

AAAY79998

ID AAAY79998 standard; peptide; 25 AA.

XX AC AAAY79998;

XX 15-MAY-2000 (first entry)

XX Optimised IgE-CH3 domain antigen peptide for human IgE.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;  
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;  
KW antibody; allergy; allergic disease; immunisation; anti-allergic;  
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Homo sapiens.

OS Synthetic.

XX WO9667293-A1.

PN





15-MAY-2000 (first entry)



CC (functional in genetically diverse subjects), in addition to a B cell  
 CC target epitope, have increased immunogenicity and may include cyclic  
 CC constraints (disulfide bridge) to stabilise conformational features and  
 CC maximize cross-reactivity to the natural target. They induce safe (non-  
 CC anaphylactogenic) antibodies. AAV7994 to AAV80084 represent amino acid  
 CC sequences used in the exemplification of the present invention  
 XX  
 SQ Sequence 25 AA;

Query Match 100.0%; Score 135; DB 3; Length 25;  
 Best Local Similarity 100.0%; Pred. No. 4.1e-14;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETKSTVSHPDLPREVRSIAKC 25  
 DB 1 CGETKSTVSHPDLPREVRSIAKC 25

RESULT 2  
 AAM50103  
 ID AAM50103 standard; protein; 424 AA.

XX AC AAM50103;

XX DT 02-SEP-2002 (first entry)

XX DE Equine IgE heavy chain constant region C-epsilon allotype A.

XX KW Cea; equine; horse; heavy chain; constant region; allotype a; IgE;  
 XX KW C-epsilon a; immunoglobulin E; antiallergic; allergy.

XX OS Equus caballus.

Key	Location/Qualifiers
FT Domain	1..97
FT Domain	/note= "CH1 domain"
FT Domain	98..205
FT Domain	/note= "CH2 domain"
FT Domain	206..312
FT Domain	/note= "CH3 domain"
FT Domain	313..424
FT Domain	/note= "CH4 domain"

XX PN WO200250280-A2.

XX PD 27-JUN-2002.

XX PF 20-DEC-2001; 2001WO-DE004810.

XX PR 21-DEC-2000; 2000DE-01064415.

XX PA (TIER-) TIERAERZTLICHE HOCHSCHULE HANNOVER.

XX PI Leibold W, Wagner B, Radbruch A;

XX DR WPI; 2002-508803/54.

XX DR N-PSDB; ABL61246.

XX PT DNA sequence encoding part of an equine immunoglobulin G, useful for  
 PT preparing isotype-specific antibodies for diagnosis and treatment of  
 PT allergy in horses.

XX PS Claim 1; Page 31-32; 37pp; German.

XX This invention describes a novel DNA (I) encoding the constant region  
 CC (CH) of the heavy chain of a horse immunoglobulin E (IgE) allotype. The  
 CC products of the invention are capable of neutralising IgE activity and  
 CC have antiallergic activity. The polynucleotides of the invention are used  
 CC to produce recombinant IgE (II), and this is used to raise specific  
 CC monoclonal antibodies (MAB). Both (II) and MAB are useful for diagnosis,  
 CC particularly of allergy in horses, and MAB can also be used for treatment  
 CC of allergies. (I) make possible production of isotype-specific  
 CC antibodies, important for precise evaluation of the immune response. This

CC sequence represents a fragment of the equine IgE heavy chain constant  
 CC region C-epsilon allotype a (Cea), used in the method of the invention  
 CC for IgE-isotype recombinant immunoglobulin production  
 XX  
 SQ Sequence 424 AA;

Query Match 83.0%; Score 112; DB 5; Length 424;  
 Best Local Similarity 95.7%; Pred. No. 5.7e-09;  
 Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GETYKSTVSHPDLPREVRSIAK 24  
 DB 288 GETYKSTVSHPDLPREVRSIAK 310

RESULT 3  
 AAM50104

ID AAM50104 standard; protein; 424 AA.

XX AC AAM50104;

XX DT 02-SEP-2002 (first entry)

XX DE Equine IgE heavy chain constant region C-epsilon allotype b.

XX KW Ceb; equine; horse; heavy chain; constant region; allotype b; IgE;  
 XX KW C-epsilon b; immunoglobulin E; antiallergic; allergy.

XX OS Equus caballus.

Key	Location/Qualifiers
FT Domain	1..97
FT Domain	/note= "CH1 domain"
FT Domain	98..205
FT Domain	/note= "CH2 domain"
FT Domain	206..312
FT Domain	/note= "CH3 domain"
FT Domain	313..424
FT Domain	/note= "CH4 domain"

XX PN WO200250280-A2.

XX PD 27-JUN-2002.

XX PF 20-DEC-2001; 2001WO-DE004810.

XX PR 21-DEC-2000; 2000DE-01064415.

XX PA (TIER-) TIERAERZTLICHE HOCHSCHULE HANNOVER.

XX PI Leibold W, Wagner B, Radbruch A;

XX DR WPI; 2002-508803/54.

XX DR N-PSDB; ABL61247.

XX PT DNA sequence encoding part of an equine immunoglobulin G, useful for  
 PT preparing isotype-specific antibodies for diagnosis and treatment of  
 PT allergy in horses.

XX PS Claim 1; Page 34-36; 37pp; German.

XX This invention describes a novel DNA (I) encoding the constant region  
 CC (CH) of the heavy chain of a horse immunoglobulin E (IgE) allotype. The  
 CC products of the invention are capable of neutralising IgE activity and  
 CC have antiallergic activity. The polynucleotides of the invention are used  
 CC to produce recombinant IgE (II), and this is used to raise specific  
 CC monoclonal antibodies (MAB). Both (II) and MAB are useful for diagnosis,  
 CC particularly of allergy in horses, and MAB can also be used for treatment  
 CC of allergies. (I) make possible production of isotype-specific  
 CC antibodies, important for precise evaluation of the immune response. This  
 CC sequence represents a fragment of the equine IgE heavy chain constant  
 CC region C-epsilon allotype b (Ceb), used in the method of the invention  
 CC for IgE-isotype recombinant immunoglobulin production

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 08:16:50 ; Search time 46.6 Seconds  
(without alignments)  
151.581 Million cell updates/sec

Title: US-09-701-623C-84  
Perfect score: 135  
Sequence: 1 CGETYSKTVSHPDLPRAVRSIAKC 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	135	100.0	25	3 AAY80077	Aay80077 Optimised
2	112	83.0	424	5 AAM50103	Aam50103 Equine Ig
3	112	83.0	424	5 AAM50104	Aam50104 Equine Ig
4	112	83.0	569	6 ABP96585	Abp96585 Horse IgE
5	105	77.8	25	3 AAY79999	Aay79999 Optimised
6	105	77.8	45	3 AAY80019	Aay80019 IGE immun
7	105	77.8	45	3 AAY80083	Aay80083 IGE immun
8	105	77.8	46	3 AAY80020	Aay80020 IGE immun
9	105	77.8	57	3 AAY80081	Aay80081 IGE immun
10	105	77.8	62	3 AAY80080	Aay80080 IGE immun
11	105	77.8	63	3 AAY80084	Aay80084 IGE immun
12	94	69.6	25	3 AAY68602	Aay68602 Peptide s
13	94	69.6	25	3 AAY91212	Aay91212 Modified
14	94	69.6	25	3 AAY79998	Aay79998 Optimised
15	94	69.6	42	3 AAY68604	Aay68604 Peptide s
16	94	69.6	42	3 AAY91216	Aay91216 Modified
17	94	69.6	42	3 AAY91215	Aay91215 Modified
18	94	69.6	42	3 AAY91217	Aay91217 Modified
19	94	69.6	42	3 AAY80014	Aay80014 IGE immun
20	94	69.6	44	7 ADD89950	Add89950 IGE pepti
21	94	69.6	45	3 AAY68605	Aay68605 Peptide s
22	94	69.6	45	3 AAY91218	Aay91218 Modified
23	94	69.6	45	3 AAY80007	Aay80007 IGE-CH3 d
24	94	69.6	45	7 ADD89951	Add89951 IGE pepti
25	94	69.6	46	3 AAY68603	Aay68603 Peptide s

26	94	69.6	46	3 AAY91213	Aay91213 Modified
27	94	69.6	46	3 AAY91214	Aay91214 Modified
28	94	69.6	46	3 AAY80011	Aay80011 IGE immun
29	94	69.6	56	3 AAY80016	Aay80016 IGE immun
30	94	69.6	59	3 AAY80010	Aay80010 IGE immun
31	94	69.6	60	3 AAY80013	Aay80013 IGE immun
32	94	69.6	60	3 AAY80015	Aay80015 IGE immun
33	94	69.6	63	3 AAY68606	Aay68606 Peptide s
34	94	69.6	63	3 AAY91219	Aay91219 Inv. epit
35	94	69.6	63	3 AAY80008	Aay80008 IGE-CH3 d
36	94	69.6	63	3 AAY80012	Aay80012 IGE immun
37	92	68.1	496	6 ABP96580	Abp96580 Cat IGE h
38	92	68.1	496	6 ABU09338	Abu09338 Feline Ig
39	92	68.1	496	6 ABU09336	Abu09336 Feline Ig
40	89	65.9	25	3 AAY80000	Aay80000 Optimised
41	89	65.9	45	3 AAY80018	Aay80018 IGE immun
42	89	65.9	46	3 AAY80017	Aay80017 IGE immun
43	87	64.4	60	3 AAY80078	Aay80078 IGE immun
44	82	60.7	114	6 ABG74772	Abg74772 Canine Ig
45	82	60.7	115	6 ABG74774	Abg74774 Human IGE

## ALIGNMENTS

RESULT 1  
AAY80077  
ID AAY80077 standard; peptide; 25 AA.  
XX  
AC AAY80077;  
XX  
DT 15-MAY-2000 (first entry)  
XX  
DE DE Optimised IGE-CH3 domain antigen peptide for horse IGE.  
XX  
KW Immunoglobulin E; IGE; epsilon heavy chain; antigenic; antigen;  
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;  
KW antibody; allergy; allergic disease; immunisation; anti-allergic;  
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.  
XX  
OS Equus caballus.  
OS Synthetic.  
XX  
PN WO9967293-A1.  
XX  
PD 29-DEC-1999.  
XX  
PF 21-JUN-1999; 99WO-US013959.  
XX  
PR 20-JUN-1998; 98US-00100287.  
XX (UNEL-) UNITED BIOMEDICAL INC.  
XX  
PI Wang CY, Walfield AM;  
XX  
DR WPI; 2000-160578/14.  
XX  
PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions  
PT for immunization against allergy.  
XX  
PS Claim 1; Page 146; 155pp; English.

The present invention describes immunoglobulin E (IGE)-CH3 domain antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic and anti-asthmatic properties. (I) induces polyclonal antibodies specific for a target effector site on the epsilon-heavy chain of IGE, and so preventing triggering and activation of mast cells and basophils and downregulation of IGE synthesis. Conjugates, or fusion peptides, containing (I) are used for active immunisation against IGE-mediated allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy dermatitis. Nucleic acids that encode these compounds are useful for recombinant production of corresponding peptides or in DNA vaccines. Conjugates of (I) that include a promiscuous T helper cell epitope

```

FT NON_TER      I
SQ SEQUENCE     478 AA; 51620 MW; 4AFCE541F3217CA1 CRC64;

Query Match           37.8%; Score 51; DB 4; Length 478;
Best Local Similarity 39.1%; Pred.No. 32;
Matches          9; Conservative    5; Mismatches   9; Indels    0; Gaps    0;

QY              2 GETYKSTVSHPDLPREVVRISIAK 24
||||:|:||||: :||
Db             324 GETTCTAAHPKLPTLTANITK 346

RESULT 13
Q72374 PRELIMINARY; PRT; 492 AA.
ID QZ374 AC QZ374;
DT DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE DE Hypothetical protein DKFP686C02218 (Fragment).
GN GN DXFP686C02218.
OS OS Homo sapiens (Human).
OC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_Taxid=9606;
RN RN [1]
RP RP SEQUENCE FROM N.A.
RC RC TISSUE=HUMAN rectum tumor;
RA RA Bloembergen H., Boecker M., Newes H.W., Weil B., Amid C., Osanger A.,
RL RLObo G., Han M., Wiemann S.;
RR Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
RW RW EMBL; BX538077; CAD98001.1; -.
KW KW Hypothetical protein.
FT FT NON_TER 1
SQ SEQUENCE 492 AA; 53776 MW; 1E7A15760FOCA74B CRC64;

Query Match           37.8%; Score 51; DB 4; Length 492;
Best Local Similarity 39.1%; Pred.No. 33;
Matches          9; Conservative    5; Mismatches   9; Indels    0; Gaps    0;

QY              2 GETYKSTVSHPDLPREVVRISIAK 24
||||:|:||||: :||
Db             338 GETTCTAAHPKLPTLTANITK 360

RESULT 14
Q99K47 PRELIMINARY; PRT; 557 AA.
ID ID Q99K47 AC Q99K47;
DT DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE DE Fibrinogen A alpha polypeptide.
GN GN PGA.
OS OS Mus musculus (Mouse);
CC CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_Taxid=10090;
RN RN [1]
RP RP SEQUENCE FROM N.A.
RA RA Strausberg R.;
RL RL Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR DR EMBL; BC005467; AAA05467.1; -.
DR DR RSSP; PG2671; 1FEA.
DR DR MGD; MGII1316726; Fga.
SQ SEQUENCE 557 AA; 61325 MW; C47FA96D1BA432DE CRC64;

Query Match           37.8%; Score 51; DB 11; Length 557;
Best Local Similarity 45.5%; Pred.No. 38;
Matches         10; Conservative    4; Mismatches   8; Indels    0; Gaps    0;

QY              1 CGETYKSTVSHPDLPREVVRISI 22
||||:|:||||: ||||: |

```



```

ID Q9CM48 PRELIMINARY; PRT; 369 AA.
AC Q9CM48: 2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE MetB.
GN METB OR PM0995.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL; AE006138; AAK03079.1; -.
DR HSP; P09355; ICS1.
DR GO; GO:0006520; P:amino acid metabolism; IEA.
DR InterPro; IPR00277; Cys Met Meta_PP.
DR Pfam; PF01053; Cys Met Meta_PP; 1.
DR PROSITE; PS00868; CYS_MET_METAB_PP; 1.
KW Complete proteome.
SQ SEQUENCE 369 AA; 40565 MW; EB8654C9FB69C5F0 CRC64;

Query Match 37.8%; Score 51; DB 16; Length 369;
Best Local Similarity 37.5%; Pred. No. 25;
Matches 9; Conservative 5; Mismatches 8; Indels 2; Gaps 1;

QY 4 TYKSTVSHPDLPRE--VVRGSIAC 25
DB 320 TYPATQTHMIPERIERARGVCNC 343

RESULT 10
Q86TT1 ID Q86TT1 PRELIMINARY; PRT; 375 AA.
AC Q86TT1: 2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Human full-length cDNA clone CSODD006YL02 of neuroblastoma of Homo sapiens (Human).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Neuroblastoma;
RA Genoscope;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Neuroblastoma;
RA Li W.B., Gruber C., Jessee J., Polayes D.;
RT "Full-length cDNA libraries and normalization.";
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX161420; CAD61894.1; -.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003597; IG_c1.
DR Pfam; PF00047; IG_3.
DR SMART; SMO0407; IGc1; 3.
DR PROSITE; PS0835; IG LIKE; 3.
DR PROSITE; PS00290; IG_MHC; 3.
KW Plasmid.
SQ SEQUENCE 375 AA; 41272 MW; 7ACD1AF4399C5E9E CRC64;

Query Match 37.8%; Score 51; DB 4; Length 375;

```

```

Best Local Similarity 39.1%; Pred. No. 25;
Matches 9; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 2 GETYKSTVSHPDLPREVVRSIK 24
DB 220 GERFCTVHTDLPSPKQIISR 242

RESULT 11
Q9NPP6 PRELIMINARY; PRT; 416 AA.
ID Q9NPP6: 2000 (Tremblrel. 15, Created)
AC Q9NPP6: 2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE Immunoglobulin heavy chain variant (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Pluvinet R., Estivill X., Escarceller M., Sumoy L.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Auffray C., Ansorge W., Ballabio A., Estivill X., Gibson K.,
RA Lehrach H., Poustka A., Lundeberg J.;
RT "The European IMAGE consortium for integrated Molecular analysis of human gene transcripts.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL389978; CAB97534.1; -.
DR HSP; P01789; IMCP.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_V.
DR Pfam; PF00047; IG_4.
DR SMART; SMO0406; IGV; 1.
DR PROSITE; PS0835; IG LIKE; 3.
DR PROSITE; PS00290; IG_MHC; 2.
FT NON TER 1
SQ SEQUENCE 416 AA; 44786 MW; 8C41708BB8A4687 CRC64;

Query Match 37.8%; Score 51; DB 4; Length 416;
Best Local Similarity 39.1%; Pred. No. 28;
Matches 9; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 2 GETYKSTVSHPDLPREVVRSIK 24
DB 262 GETFTCTAHPHLPKTPITANITK 284

RESULT 12
Q7Z379 PRELIMINARY; PRT; 478 AA.
ID Q7Z379: 2003 (Tremblrel. 25, Created)
AC Q7Z379: 2003 (Tremblrel. 25, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hypothetical protein DKFZp686K04218 (Fragment).
GN DKFZP686K04218.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human rectum tumor;
RA Bloecker H., Boeche M., Mewes H.W., Weill B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX538066; CAD97996.1; -.
KW Hypothetical protein.

```



```

DR PROSITE, PS00290; IG_MHC; 3.
KW Receptor; Signal. 18 POTENTIAL.
FT CHAIN 19 684 NOVEL ANTIGEN RECEPTOR..
SQ SEQUENCE 684 AA; 75224 MW; 2FF9D2071CDA6DFD CRC64;

Query Match 39.3%; Score 53; DB 13; Length 684;
Best Local Similarity 47.6%; Pred. No. 24;
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 4 TYKSTVSHPDLPREVRSIAK 24
DB 642 TYSLVGHFSLNRDLIRSTNK 662
||| ||| ||| ||| ||| ||| |||
||| ||| ||| ||| ||| ||| |||

RESULT 4
Q7VHN6 PRELIMINARY; PRT; 209 AA.
AC Q7VHN6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
DE GN HH0929.
OS Helicobacter hepaticus.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=32025;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 51449 / 3B1;
RX MEDLINE=22709201; PubMed=12810954;
RA Suerbaum S., Josenhans C., Sterzenbach T., Drescher B., Brandt P.,
RA Bell M., Droege M., Fartmann B., Fischer H.-P., Ge Z., Hoerster A.,
RA Holland R., Klein K., Koenig J., Macko L., Mendz G.L., Nyakatura G.,
RA Schauer D.B., Shen Z., Weber J., Frosch M., Fox J.G.;
RT "The complete genome sequence of the carcinogenic bacterium
RT Helicobacter hepaticus."
RL Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906(2003).
RW EMBL; AF017146; AAP7526.1; .
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 209 AA; 23936 MW; ED0B50093E84F6FA CRC64;

Query Match 38.5%; Score 52; DB 16; Length 209;
Best Local Similarity 27.3%; Pred. No. 9.4;
Matches 9; Conservative 7; Mismatches 9; Indels 8; Gaps 1;

QY 1 CGEYKSTVSHPDLPF-----EVVRSIAKC 25
DB 20 CGEYQNTTFHFPSPQAALENNLESIIKNCSLC 52
||| ||| ||| ||| ||| ||| |||
||| ||| ||| ||| ||| ||| |||

RESULT 5
O17620 PRELIMINARY; PRT; 394 AA.
ID O17620;
AC O17620;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE C29F7.1 protein.
DE GN C29F7.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA McMurry A.A.
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 08:18:11 ; Search time 31.4 Seconds  
(without alignments)  
251.209 Million cell updates/sec

Title: US-09-701-623C-84  
Perfect score: 135  
Sequence: 1 CGEYKSTVSHPDLPREVRSIAC 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL 25.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mbc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54	40.0	137	Q57464	Q57464 photobacter
2	54	40.0	599	Q7VL73	Q7VL73 haemophilus
3	53	39.3	684	Q90544	Q90544 ginglymosto
4	52	38.5	209	Q7VH6	Q7VH6 helicobacte
5	52	38.5	394	Q17620	Q17620 caenorhabdi
6	52	38.5	1093	Q8SD16	Q8SD16 pseudomonas
7	51	37.8	93	Q31260	Q31260 rattus norv
8	51	37.8	341	Q08284	Q08284 mus musculu
9	51	37.8	369	Q9CM48	Q9CM48 pasteurilla
10	51	37.8	375	Q86T11	Q86T11 homo sapien
11	51	37.8	416	Q9NPP6	Q9NPP6 homo sapien
12	51	37.8	478	Q7Z379	Q7Z379 homo sapien
13	51	37.8	492	Q7Z374	Q7Z374 homo sapien
14	51	37.8	557	Q99K47	Q99K47 mus musculu
15	51	37.8	588	Q8WUX4	Q8WUX4 homo sapien
16	51	37.8	597	Q9BU10	Q9BU10 homo sapien

17	51	37.8	597	4	Q9BQB8	Q9BQB8 homo sapien
18	51	37.8	597	4	Q9BBB9	Q9BBB9 homo sapien
19	51	37.8	613	4	Q9EY0	Q9EY0 homo sapien
20	51	37.8	613	4	Q8WUK1	Q8WUK1 homo sapien
21	51	37.8	614	4	Q9GGA6	Q9GGA6 homo sapien
22	51	37.8	618	4	Q9GAA6	Q9GAA6 homo sapien
23	51	37.8	882	5	Q9NE67	Q9NE67 leishmania
24	50.5	37.4	893	5	Q02099	Q02099 caenorhabdi
25	50	37.0	245	2	Q9EV20	Q9EV20 listeria mo
26	49	36.3	278	16	Q8D5F0	Q8D5F0 vibrio vuln
27	49	36.3	364	2	Q8GHJ5	Q8GHJ5 streptococc
28	49	36.3	426	2	Q8GDR1	Q8GDR1 heliobacill
29	49	36.3	473	11	Q91Z05	Q91Z05 mus musculu
30	49	36.3	474	11	Q9R3H6	Q9R3H6 mus musculu
31	49	36.3	520	10	Q48786	Q48786 arabidopsis
32	49	36.3	712	5	Q960E2	Q960E2 drosophila
33	49	36.3	1000	5	Q8GNQ0	Q8GNQ0 drosophila
34	49	36.3	1105	5	Q9VN44	Q9VN44 drosophila
35	48.5	35.9	125	11	Q8BW77	Q8BW77 mus musculu
36	48.5	35.9	255	11	Q8BUP3	Q8BUP3 mus musculu
37	48.5	35.9	275	11	Q8BXL3	Q8BXL3 mus musculu
38	48.5	35.9	279	11	Q9D9F8	Q9D9F8 mus musculu
39	48	35.6	113	2	Q8KK57	Q8KK57 proteus vul
40	48	35.6	183	16	Q89W99	Q89W99 bradyrhizob
41	48	35.6	326	5	Q9VIJ0	Q9VIJ0 drosophila
42	48	35.6	335	11	Q88283	Q88283 mesocricetu
43	48	35.6	364	5	Q8SZU0	Q8SZU0 drosophila
44	48	35.6	425	11	Q9DBZ2	Q9DBZ2 mus musculu
45	48	35.6	532	12	Q8JNC8	Q8JNC8 macropodid

ALIGNMENTS

RESULT 1  
Q57464 PRELIMINARY; PRT; 137 AA.  
ID Q57464  
AC Q57464;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE RPOE, LEPA and ORF 2, 3, 4 genes (Fragment).  
GN LEPA.  
OS Photobacterium profundum (Photobacterium sp. (strain SS9)).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; Photobacterium.  
OX NCBI\_TaxID=74109;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SS9;  
RA Chi E., Bartlett D.H.;  
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SS9;  
RA Chi E., Bartlett D.H.;  
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL; L41688; AAA65230.1; -;  
DR EMBL; L41667; AAB31927.1; -;  
DR PIR; S70214; S70214.  
DR HSP; P13551; IDAR.  
DR GO; GO:000525; F:GTP binding; IEA.  
DR GO; GO:0003746; F:translation elongation factor activity; IEA.  
DR GO; GO:0006414; P:translational elongation; IEA.  
DR InterPro; IPR000795; EF\_GTPbind.  
DR Pfam; PF00009; GTP\_EFTU; 1.  
DR PRINTS; PR00315; EFONGATNECT.  
DR PROSITE; PS00301; EFACITOR\_GTP; 1.  
KW GTP-binding; Protein biosynthesis.  
FT NON TER 137  
SQ SEQUENCE 137 AA; 15016 MW; 7FC5E05FDEAF1F3E CRC64;

Query Match 40.0%; Score 54; DB 2; Length 137;

RA Kerr M.A.;  
RT "The structure and function of human IgA."  
RL Biochem. J. 271:285-296(1990).  
CC -!- FUNCTION: Ig alpha is the major immunoglobulin class in body  
CC secretions. It may serve both to defend against local infection  
CC and to prevent access of foreign antigens to the general  
CC immunologic system.  
CC -!- SUBUNIT: Monomeric or polymeric.  
CC -!- MISCELLANEOUS: The sequence of the A2m(1) allotype is shown.  
CC -!- SIMILARITY: Contains 3 immunoglobulin-like domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; J00221; AAB59396.1; ALT\_INIT.  
DR PIR; B22360; B22360.  
DR HSSP; P01810; 2PBJ.  
DR Genew; HGNC:5479; IGHA2.  
DR MIN; I47000; -.  
DR GO; GO:0005624; C:membrane fraction; NAS.  
DR GO; GO:0003823; F:antigen binding; TAS.  
DR GO; GO:0006955; P:immune response; NAS.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003597; Ig cl.  
DR InterPro; IPR003006; Ig\_MHC.  
DR Pfam; PF00047; Ig; 3.  
DR SMART; SM00407; IGcl; 2.  
DR PROSITE; PS00835; IG LIKE; 3.  
DR PROSITE; PS00290; IG\_MHC; 2.  
DR KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein; Repeat.  
FT NON TER 1 1  
FT DOMAIN 6 98 IG-LIKE 1.  
FT DOMAIN 112 207 IG-LIKE 2.  
FT DOMAIN 215 317 IG-LIKE 3.  
FT DISULFID 26 85 PROBABLE.  
FT DISULFID 101 101 INTERCHAIN (WITH LIGHT CHAIN) (PROBABLE).  
FT DISULFID 109 109 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).  
FT DISULFID 110 167 PROBABLE.  
FT DISULFID 134 191 PROBABLE.  
FT DISULFID 169 169 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).  
FT DISULFID 179 179 INTERCHAIN (WITH HEAVY CHAIN IN ANOTHER  
FT SUBUNIT) (POTENTIAL).  
FT DISULFID 237 300 PROBABLE.  
FT DISULFID 339 339 INTERCHAIN (WITH J CHAIN) (PROBABLE).  
FT CARBOHYD 47 47 N-LINKED (GLCNAC. .).  
FT CARBOHYD 92 92 N-LINKED (GLCNAC. .).  
FT CARBOHYD 131 131 N-LINKED (GLCNAC. .).  
FT CARBOHYD 205 205 N-LINKED (GLCNAC. .).  
FT CARBOHYD 327 327 N-LINKED (GLCNAC. .).  
FT VARIANT 93 93 P -> S (IN A2M(2) ALLOTYP).  
FT VARIANT 102 102 /FTid=VAR 003879.  
FT VARIANT 279 279 P -> R (IN A2M(2) ALLOTYP).  
FT VARIANT 296 296 /FTid=VAR 003880.  
FT VARIANT 326 326 F -> Y (IN A2M(2) ALLOTYP).  
FT VARIANT 335 335 D -> E (IN A2M(2) ALLOTYP).  
FT VARIANT 340 AA; 36508 MW; 98922700756F3276 CRC64;  
FT SEQUENCE 340 AA; 36508 MW; 98922700756F3276 CRC64;  
SQ  
Query Match 37.8%; Score 51; DB 1; Length 340;  
Best Local Similarity 39.1%; Pred. No. 3.6;  
Matches 9; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 2-GETYKSTVSHPDLPREVRSIAK 24

Db 186 GETFTCTAAHPKELKPLTANITK 208

Search completed: February 26, 2004, 08:24:23  
Job time : 6.4 secs

Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 -!- SIMILARITY: Belongs to the GTP-binding elongation factor family.  
 LepA subfamily.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; AE005781; AAK23018.1; ALT\_INIT.  
 HSSP; F13551; IELO.  
 TIGR; CC1034; -.  
 HAMAP; MF\_00071; -; 1.  
 InterPro; IPR000795; EF\_GTPbind.  
 InterPro; IPR000640; EFG\_C.  
 InterPro; IPR009022; EFG\_III\_V.  
 InterPro; IPR004161; EFTU\_D2.  
 InterPro; IPR006297; LepA.  
 InterPro; IPR005225; Small\_GTP.  
 InterPro; IPR009000; Translat\_factor.  
 Pfam; PF00679; EFG\_C; 1.  
 Pfam; PF03144; GTP\_EFTU\_D2; 1.  
 PRINTS; PR00315; ELONGATNFCT.  
 TIGRFAMS; TIGR01393; lepa; 1.  
 TIGRFAMS; TIGR00231; small\_GTP; 1.  
 PROSITE; PS00301; EFATOR\_GTP; 1.  
 GTP-binding; Complete proteome.  
 NP\_BIND 20 27 GTP (BY SIMILARITY).  
 NP\_BIND 86 90 GTP (BY SIMILARITY).  
 NP\_BIND 140 143 GTP (BY SIMILARITY).  
 SEQUENCE 606 AA; 66856 MW; 94FE045B04EEF615 CRC64;  
 -----  
 Query Match 38.5%; Score 52; DB 1; Length 606;  
 Best Local Similarity 46.4%; Pred. No. 4.7;  
 Matches 13; Conservative 1; Mismatches 10; Indels 4; Gaps 1;  
 -----  
 QY 2 GETY-----KSTVSHPDLPREVVRSIAC 25  
 DB 77 GETILMDTPGHVDFAYEVSRLAAC 104  
 -----  
 RESULT 14  
 LEPA STAEP STANDARD; PRT; 607 AA.  
 AC Q8CPL3;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE GTP-Binding protein lepa.  
 GN LEPA OR S1271.  
 OS Staphylococcus epidermidis.  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 OX NCBI\_TaxID=1282;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 12228;  
 RX PubMed=12950922;  
 RA Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,  
 RA Qin Z.-Q., Mao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,  
 RA Yuan Z.-H., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.;  
 RT "Genome-based analysis of virulence genes in a non-biofilm-forming  
 RT Staphylococcus epidermidis strain (ATCC 12228).";  
 RL Mol. Microbiol. 49:1577-1593(2003).  
 CC -!- SIMILARITY: Belongs to the GTP-binding elongation factor family.  
 CC LepA subfamily.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; AE016748; AAO04870.1; -.  
 HAMAP; MF\_00071; -; 1.  
 InterPro; IPR000795; EF\_GTPbind.  
 InterPro; IPR000640; EFG\_C.  
 InterPro; IPR009022; EFG\_III\_V.  
 InterPro; IPR004161; EFTU\_D2.  
 InterPro; IPR006297; LepA.  
 InterPro; IPR005225; Small\_GTP.  
 InterPro; IPR009000; Translat\_factor.  
 Pfam; PF00679; EFG\_C; 1.  
 Pfam; PF03144; GTP\_EFTU\_D2; 1.  
 PRINTS; PR00315; ELONGATNFCT.  
 TIGRFAMS; TIGR01393; lepa; 1.  
 TIGRFAMS; TIGR00231; small\_GTP; 1.  
 PROSITE; PS00301; EFATOR\_GTP; 1.  
 GTP-binding; Complete proteome.  
 NP\_BIND 20 27 GTP (BY SIMILARITY).  
 NP\_BIND 86 90 GTP (BY SIMILARITY).  
 NP\_BIND 140 143 GTP (BY SIMILARITY).  
 SEQUENCE 607 AA; 68275 MW; DA714C861EBC589F CRC64;  
 -----  
 Query Match 38.5%; Score 52; DB 1; Length 607;  
 Best Local Similarity 46.4%; Pred. No. 4.7;  
 Matches 13; Conservative 1; Mismatches 10; Indels 4; Gaps 1;  
 -----  
 QY 2 GETYK-----STVSHPDLPREVVRSIAC 25  
 DB 77 GETYFHLIDTPGHVDFYEVSRSLAAC 104  
 -----  
 RESULT 15  
 ALC2 HUMAN STANDARD; PRT; 340 AA.  
 ID ALC2 HUMAN  
 AC P01877;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Ig alpha-2 chain C region.  
 GN IGHA2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=84130179; PubMed=6421489;  
 RA Flanagan J.G., Lefranc M.-P., Rabbitts T.H.;  
 RT "Mechanisms of divergence and convergence of the human immunoglobulin  
 RT alpha 1 and alpha 2 constant region gene sequences.";  
 RL Cell 36:681-688(1984).  
 RN [2]  
 RP SEQUENCE (BUT).  
 RX MEDLINE=78137069; PubMed=416441;  
 RA Torano A., Putnam F.W.;  
 RT "Complete amino acid sequence of the alpha 2 heavy chain of a human  
 RT IgA2 immunoglobulin of the A2m (2) allotype.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 75:966-969(1978).  
 RN [3]  
 RP SEQUENCE (MYELOMA PROTEIN LAN).  
 RX MEDLINE=79180140; PubMed=286295;  
 RA Tsukida Y., Wang C.-C., Putnam F.W.;  
 RT "Structure of the A2m(1) allotype of human IgA -- a recombinant  
 RT molecule.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 76:1104-1108(1979).  
 RN [4]  
 RP REVIEW.  
 RX MEDLINE=91054387; PubMed=2241915;

```
Best Local Similarity 46.4%; Pred. No. 2.3;
Matches 13; Conservative 2; Mismatches 9; Indels 4; Gaps 1;

QY 2 GETYK----STVSHDLPREVVRISIAK 25
DQ 72 GETYELNIDTPGHVDFSEVSRSLAAC 99

RESULT 11
LEPA_CHLTR
ID LEPA_CHLTR STANDARD; PRT; 602 AA.
AC 084067;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE GTP-binding protein lepa.
GN LEPA OR CT064.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/UW-3/Cx;
RX MEDLINE=99000809; PubMed=9784136;
RA Stephens R.S., Kallman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis.";
RL Science 282:754-759(1998).
CC -!- SIMILARITY: Belongs to the GTP-binding elongation factor family.
CC LePA subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; AE001281; AAC67655.1; -.
CC F1; B71561; B71561.
CC HSP; P13551; LELO.
CC HANAP; MF 00071; -.
CC InterPro; IPR000795; EF_GTPbind.
CC InterPro; IPR000640; EFG_C.
CC InterPro; IPR009022; EFG_III_V.
CC InterPro; IPR004161; EFTU_D2.
CC InterPro; IPR006297; LePA.
CC InterPro; IPR005225; Small_GTP.
CC InterPro; IPR009000; Translat_factor.
CC Pfam; PF006679; EFG_C; 1.
CC Pfam; PF03144; GTP_EFTU_D2; 1.
CC PRINTS; PR00315; ELONGATNFT.
CC TIGRfams; TIGR01393; lepa; 1.
CC TIGRfams; TIGR00231; small_GTP; 1.
CC PROSITE; PS00301; EFACTOR_GTP; FALSE_NEG.
KW GTP-binding; Complete proteome.
FT NP_BIND 16 23 GTP (BY SIMILARITY).
FT NP_BIND 81 85 GTP (BY SIMILARITY).
FT NP_BIND 135 138 GTP (BY SIMILARITY).
SQ SEQUENCE 602 AA; 67435 MW; 98F67C990196EE4 CRC64;

Query Match 40.0%; Score 54; DB 1; Length 602;
Best Local Similarity 46.4%; Pred. No. 2.3;
Matches 13; Conservative 2; Mismatches 9; Indels 4; Gaps 1;

QY 2 GETYK----STVSHDLPREVVRISIAK 25
DQ 72 GETYELNIDTPGHVDFSEVSRSLAAC 99
```

```
RESULT 12
MUC_CANFA
ID MUC_CANFA STANDARD; PRT; 450 AA.
AC P01874;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig mu chain C region.
DE Canis familiaris (Dog).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE OF 1-177 (MOO).
RX MEDLINE=80077682; PubMed=117299;
RA McCumber L.J., Capra J.D.;
RT "The complete amino-acid sequence of a canine mu chain.";
RL Mol. Immunol. 16:565-570(1979).
RN [2]
RP SEQUENCE OF 178-450 (MOO).
RX MEDLINE=78180587; PubMed=653360;
RA Wasserman R.L., Capra J.D.;
RT "Amino acid sequence of the Fc region of a canine immunoglobulin M:
RT interspecies homology for the IgM class.";
RL Science 200:1159-1161(1978).
DR PIR; A93131; MHDG.
DR HSSP; P01857; LFCL.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF000407; Ig; 4.
DR SMART; SM00407; IGc1; 2.
DR PROSITE; PS00835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
SQ SEQUENCE 450 AA; 48895 MW; 9D460DA9D1012F5D CRC64;

Query Match 38.9%; Score 52.5; DB 1; Length 450;
Best Local Similarity 50.0%; Pred. No. 2.9;
Matches 12; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

QY 2 GETYKTVSHDLPREVVRISIAK 24
DQ 405 GETYTCVAHESLPNRTSVRSVDK 428

RESULT 13
LEPA_CAUCR
ID LEPA_CAUCR STANDARD; PRT; 606 AA.
AC Q9A9F4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE GTP-Binding protein lepa.
GN LEPA OR CC1034.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Risen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Ullrich J., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
```







```

DR TIGR; HI0016; -.
DR HAMAP; MF_00071; -.
DR InterPro; IPR000795; EF GTPbind.
DR InterPro; IPR000640; EFG C.
DR InterPro; IPR000922; EFG III V.
DR InterPro; IPR004161; EFTU D2.
DR InterPro; IPR006297; LepA.
DR InterPro; IPR005225; Small GTP.
DR InterPro; IPR009000; Translat_factor.
DR Pfam; PF00679; EFG_C; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR TIGRfams; TIGR01393; lepa; 1.
DR TIGRfams; TIGR00231; small GTP; 1.
DR PROSITE; PS00301; EFATOR_GTP; 1.
KW Membrane; GTP-binding; Complete proteome.
FT NP_BIND 11 18 GTP (BY SIMILARITY).
FT NP_BIND 77 81 GTP (BY SIMILARITY).
FT NP_BIND 131 134 GTP (BY SIMILARITY).
SQ SEQUENCE 598 AA; 66333 MW; FDDBIAC3516BF11E CRC64;

Query Match 40.0%; Score 54; DB 1; Length 598;
Best Local Similarity 46.4%; Pred.No. 2.3;
Matches 13; Conservative 2; Mismatches 9; Indels 4; Gaps 1;

QY 2 GETYK----STVSHDPLPREVVRGIAC 25
      |||||:  |||||:
DB 68 GETYQLNFIDTPGHVDFSVSRSLAAC 95

RESULT 7
LEPA_PASMU
ID LEPA_PASMU STANDARD; PRT; 598 AA.
AC P57806;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE GTP-Binding protein lepa.
GN LEPA OR PM0063.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RL "Complete genomic sequence of Pasteurella multocida pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -!- SIMILARITY: Belongs to the GTP-binding elongation factor family.
CC LepA subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE006042; AAK02147.1; -.
DR HSP; P13551; 1DAR.
DR HAMAP; MF_00071; -.
DR InterPro; IPR000795; EF GTPbind.
DR InterPro; IPR000640; EFG C.
DR InterPro; IPR009022; EFG III V.
DR InterPro; IPR004161; EFTU D2.
DR InterPro; IPR006297; LepA.
DR InterPro; IPR005225; Small GTP.
DR InterPro; IPR009000; Translat_factor.
DR Pfam; PF00679; EFG_C; 1.

```

```

DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR TIGRfams; TIGR01393; lepa; 1.
DR TIGRfams; TIGR00231; small GTP; 1.
DR PROSITE; PS00301; EFATOR_GTP; 1.
KW GTP-binding; Complete proteome.
FT NP_BIND 11 18 GTP (BY SIMILARITY).
FT NP_BIND 77 81 GTP (BY SIMILARITY).
FT NP_BIND 131 134 GTP (BY SIMILARITY).
SQ SEQUENCE 598 AA; 66211 MW; 4D41723C2AEBFD10 CRC64;

Query Match 40.0%; Score 54; DB 1; Length 598;
Best Local Similarity 46.4%; Pred.No. 2.3;
Matches 13; Conservative 2; Mismatches 9; Indels 4; Gaps 1;

QY 2 GETYK----STVSHDPLPREVVRGIAC 25
      |||||:  |||||:
DB 68 GETYQLNFIDTPGHVDFSVSRSLAAC 95

RESULT 8
LEPA_ECOLI
ID LEPA_ECOLI STANDARD; PRT; 599 AA.
AC P07682; P76590;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE GTP-binding protein lepa.
GN LEPA OR B2569 OR C3093 OR Z3851 OR ECS3435 OR SF2631 OR S2804.
OS Escherichia coli,
OS Escherichia coli O6,
OS Escherichia coli O157:H7, and
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 217992, 83334, 623;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12;
RX MEDLINE=85207751; PubMed=2987248;
RA March P.E., Inouye M.;
RT "Characterization of the lep operon of Escherichia coli.
RT Identification of the promoter and the gene upstream of the signal
RT Peptidase I gene.";
RL J. Biol. Chem. 260:7206-7213(1985).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12;
RA Nashimoto H., Saito N.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / MG1655;
RX MEDLINE=9742617; PubMed=9778503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
RN [5]

```

```

RL Lancelot 361:743-749(2003).
CC -!- SIMILARITY: Belongs to the GTP-binding elongation factor family.
CC Lepa subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP005082; BAC60837.1; -.
DR HAVAP; MF 00071; -.
DR InterPro; IPR000795; EF_GTPbind.
DR InterPro; IPR000640; EFG_C.
DR InterPro; IPR006297; LepA.
DR InterPro; IPR009022; EFG_III_V.
DR InterPro; IPR004161; EFTU_D2.
DR InterPro; IPR009000; Translat_factor.
DR Pfam; PF00679; EFG_C; 1.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR PROSITE; PS00301; EFATOR_GTP; 1.
DR GTP-binding; Complete proteome.
FT NP_BIND 11 18 GTP (BY SIMILARITY).
FT NP_BIND 77 81 GTP (BY SIMILARITY).
FT NP_BIND 131 134 GTP (BY SIMILARITY).
SQ SEQUENCE 597 AA; 65914 MW; 34E0EDBF1FCB864 CRC64;

Query Match 40.0%; Score 54; DB 1; Length 597;
Best Local Similarity 46.4%; Pred. No. 2.3;
Matches 13; Conservative 2; Mismatches 9; Indels 4; Gaps 1;

QY 2 GETYK----STVSHDPLPREVVRGIAC 25
| | | | |
DB 68 GETYQLNFIDTPGHVDFAVEYSRSLAAC 95

RESULT 5
LEPA_VIBVU
ID LEPA_VIBVU STANDARD; PRT; 597 AA.
AC Q8DC78;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE GTP-Binding protein lepa.
GN LEPA OR VV11563.
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the GTP-binding elongation factor family.
CC Lepa subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE016802; AAC09987.1; -.
DR HAVAP; MF 00071; -.
DR InterPro; IPR000795; EF_GTPbind.

```

```

DR InterPro; IPR000640; EFG_C.
DR InterPro; IPR009022; EFG_III_V.
DR InterPro; IPR004161; EFTU_D2.
DR InterPro; IPR006297; LepA.
DR InterPro; IPR005225; Small_GTP.
DR InterPro; IPR009000; Translat_factor.
DR Pfam; PF00679; EFG_C; 1.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR PROSITE; PS00301; EFATOR_GTP; 1.
DR TIGRFAMs; TIGR01393; lepa; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
DR PROSITE; PS00301; EFATOR_GTP; 1.
DR GTP-binding; Complete proteome.
KW FT NP_BIND 11 18 GTP (BY SIMILARITY).
FT NP_BIND 77 81 GTP (BY SIMILARITY).
FT NP_BIND 131 134 GTP (BY SIMILARITY).
SQ SEQUENCE 597 AA; 65961 MW; 5845994FF33CR245 CRC64;

Query Match 40.0%; Score 54; DB 1; Length 597;
Best Local Similarity 46.4%; Pred. No. 2.3;
Matches 13; Conservative 2; Mismatches 9; Indels 4; Gaps 1;

QY 2 GETYK----STVSHDPLPREVVRGIAC 25
| | | | |
DB 68 GETYQLNFIDTPGHVDFAVEYSRSLAAC 95

RESULT 6
LEPA_HAEIN
ID LEPA_HAEIN STANDARD; PRT; 598 AA.
AC F43729;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE GTP-binding protein lepa.
GN LEPA OR HI0016.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RC MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uitterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.D., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RA "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
CC -!- SUBCELLULAR LOCATION: PRESENT IN THE CYTOPLASMIC MEMBRANE AND IS
CC ALSO FOUND IN THE PERIPLASM (BY SIMILARITY).
CC -!- SIMILARITY: Belongs to the GTP-binding elongation factor family.
CC Lepa subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U32687; AAC21694.1; -.
DR PIR; I64042; I64042.
DR HSSP; P13551; 1DAR.

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
CC EMBL; X01857; CAA25977.1; --  
CC EMBL; X01857; CAA25978.1; --  
CC PIR; A02144; EHMS.  
CC PIR; A02145; EHMS.  
CC HSSP; P01854; 1IGE.  
CC InterPro; IPR007110; Ig-Like.  
CC InterPro; IPR003597; Ig\_C1.  
CC InterPro; IPR003006; Ig\_MHC.  
CC Pfam; PF00047; Ig; 4.  
CC SMART; SM00407; IGC1; 2.  
CC PROSITE; PS00835; IG-Like; 4.  
CC PROSITE; PS00290; IG\_MHC; 3.  
CC Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.  
CC NON\_TER 1 1  
CC DOMAIN 1 90 CH1.  
CC DOMAIN 91 197 CH2.  
CC DOMAIN 198 304 CH3.  
CC DOMAIN 305 421 CH4.  
CC BY SIMILARITY.  
CC DISULFID 23 75  
CC DISULFID 121 180  
CC DISULFID 226 285  
CC DISULFID 330 392  
CC BY SIMILARITY.  
CC N-LINKED (GLCNAC. .) (POTENTIAL).  
CC CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC CARBOHYD 84 84 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC CARBOHYD 95 95 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC CARBOHYD 166 166 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC CARBOHYD 238 238 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC CARBOHYD 261 261 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC CARBOHYD 365 365 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC CARBOHYD 415 415 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC SEQUENCE 421 AA; 47320 MW; 8F909B1F30A06B47 CRC64;

Query Match 48.9%; Score 66; DB 1; Length 421;  
Best Local Similarity 52.2%; Pred. No. 0.025;  
Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 GETYKSTVSHDPLPREVVRISIAK 24  
DB 280 GYGQICVDHDFPKPIVRSITK 302  
|||:|||||:|||||

RESULT 3  
EPC\_RAT  
ID EPC\_RAT STANDARD; PRT; 429 AA.  
AC P01855;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Ig epsilon chain C region.  
OS Rattus norvegicus (Rat).  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OC NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A. (IMMUNOCYTOMA IR2).  
RC STRAIN=LOU/C/WSL;  
RX MEDLINE=83064537; PubMed=6292865;  
RA Hellman L., Pettersson U., Engstrom A., Karlsson T., Bennich H.;  
RT "Structure and evolution of the heavy chain from rat immunoglobulin  
RT E.";  
RL Nucleic Acids Res. 10:6041-6049(1982).  
RN [2]  
RP SEQUENCE OF 168-342 FROM N.A. (MYELOMA IR162).

RX MEDLINE=83182019; PubMed=6820340;  
RA Kinsvogel W.R., Reddy E.P., Moore J.M., Faust C.H. Jr.;  
RT "A cloned cDNA probe for rat immunoglobulin epsilon heavy chain:  
RT construction, identification, and DNA sequence.";  
RL DNA 1:335-343(1982).  
RN [3]  
RP SEQUENCE OF 205-306 FROM N.A.  
RX MEDLINE=82174576; PubMed=6803238;  
RA Hellman L., Pettersson U., Bennich H.;  
RT "Characterization and molecular cloning of the mRNA for the heavy  
RT (epsilon) chain of rat immunoglobulin E.";  
RL Proc. Natl. Acad. Sci. U.S.A. 79:1264-1268(1982).  
CC !- SIMILARITY: Contains 4 immunoglobulin-like domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; J00744; AAA41379.1; ALT\_INIT.

DR PIR; A93442; EHRT.

DR HSSP; P01854; 1IGE.

DR InterPro; IPR007110; Ig-Like.

DR InterPro; IPR003597; Ig\_C1.

DR InterPro; IPR003006; Ig\_MHC.

DR Pfam; PF00047; Ig; 4.

DR SMART; SM00407; IGC1; 1.

DR PROSITE; PS00835; IG-Like; 4.

DR PROSITE; PS00290; IG\_MHC; 3.

DR Immunoglobulin domain; Immunoglobulin C region; Repeat.

FT NON\_TER 1 1

FT DOMAIN 6 89 IG-Like 1.

FT DOMAIN 103 201 IG-Like 2.

FT DOMAIN 205 305 IG-Like 3.

FT DOMAIN 314 414 IG-Like 4.

FT DOMAIN 168 168 R -> N (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

FT CONFLICT 308 308 P -> L (IN REF. 2).

Query Match 48.9%; Score 66; DB 1; Length 429;  
Best Local Similarity 52.2%; Pred. No. 0.025;  
Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 GETYKSTVSHDPLPREVVRISIAK 24  
DB 284 GEGYQCRVDHDFPKPIVRSITK 306  
|||:|||||:|||||

RESULT 4  
LEPA\_VIBPA  
ID LEPA\_VIBPA STANDARD; PRT; 597 AA.  
AC Q87LN7;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE GTP-Binding protein lepa.  
GN LEPA OR VP2574.  
OS Vibrio parahaemolyticus.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; Vibrio.  
OC NCBI\_TaxID=670;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RMD 2210633 / Serotype O3:K6;  
RX MEDLINE=22508454; PubMed=12620739;  
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,  
RA Iijima Y., Nishijima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,  
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;  
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism  
RT distinct from that of V. cholerae.";

```
RP 3D-STRUCTURE MODELING.
RX MEDLINE=87089848; PubMed=3796618;
RA Padlan E.A.; Davies D.R.;
RT "A model of the Fc of immunoglobulin E.";
RL Mol. Immunol. 23:1063-1075(1986).
CC -1- SIMILARITY: Contains 4 immunoglobulin-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L00022; AAB59424.1; ALT_INIT.
DR PIR; A22771; EHHU.
DR PDB; 1IGE; 15-JUL-92.
DR PDB; 1FP5; 30-JAN-02.
DR PDB; 1G84; 16-MAY-01.
DR PDB; 1OOV; 18-SEP-02.
DR Genew; HGNC:5522; IGHE.
DR MIN; 147180; -.
DR GO; GO:0003823; F:antigen binding; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; IG-Like.
DR InterPro; IPR003597; IG-G1.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00407; IGc1; 4.
DR PROSITE; PS00835; IG LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 3.
DR Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW 3D-structure; Repeat. 1
FT NON TER 1
FT DOMAIN 6 103 IG-LIKE 1.
FT DOMAIN 112 210 IG-LIKE 2.
FT DOMAIN 214 318 IG-LIKE 3.
FT DOMAIN 324 423 IG-LIKE 4.
FT DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 15 105
FT DISULFID 29 85
FT DISULFID 121 121 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 135 193 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 209 299
FT DISULFID 239 299
FT DISULFID 345 405
FT CARBOHYD 21 21 N-LINKED (GLCNAC. . .).
FT CARBOHYD 49 49 N-LINKED (GLCNAC. . .).
FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .).
FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .).
FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .).
FT CARBOHYD 275 275 N-LINKED (GLCNAC. . .).
FT VARIANT 359 359 W -> L (POSSIBLE POLYMORPHISM).
FT STRAND 110 110 /FTid=VAR_003885.
FT STRAND 113 117
FT HELIX 122 124
FT TURN 125 125
FT STRAND 130 141
FT STRAND 146 150
FT STRAND 155 155
FT STRAND 159 161
FT STRAND 165 166
FT TURN 168 169
FT STRAND 172 181
FT HELIX 182 186
FT TURN 187 188
FT STRAND 192 196
FT TURN 198 199
FT TURN 201 202
FT STRAND 205 207
FT STRAND 218 222
```

```
FT TURN 226 229
FT TURN 231 232
FT STRAND 236 242
FT STRAND 252 252
FT STRAND 255 257
FT TURN 258 259
FT STRAND 260 260
FT TURN 265 266
FT STRAND 271 274
FT TURN 275 276
FT STRAND 277 285
FT HELIX 288 293
FT TURN 294 294
FT STRAND 297 302
FT TURN 304 305
FT STRAND 310 314
FT STRAND 322 322
FT STRAND 325 329
FT HELIX 334 337
FT STRAND 340 348
FT STRAND 350 351
FT STRAND 356 360
FT TURN 366 369
FT STRAND 374 374
FT STRAND 377 378
FT TURN 380 381
FT STRAND 384 385
FT STRAND 387 393
FT HELIX 394 398
FT TURN 399 400
FT STRAND 404 408
FT STRAND 418 421
FT SEQUENCE 428 AA; 47019 MW; 25C4CA072AA558A0 CRC64;

Query Match 52.6%; Score 71; DB 1; Length 428;
Best Local Similarity 56.5%; Pred. No. 0.0045; 6; Indels 0; Gaps 0;
Matches 13; Conservative 4; Mismatches 6;

QY 2 GETYKSTVSHPDLPREVVRSTAK 24
   ||||:|:|:|:|:|:|:|:|
Db 294 GETYQCRVTHPLPALMRSTTK 316

RESULT 2
EPC_MOUSE STANDARD; PRT; 421 AA.
ID EPC_MOUSE
AC P06336; P01856;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig epsilon chain C region.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84236092; PubMed=6329728;
RA Ishida N., Ueda S., Hayashida H., Miyata T., Honjo T.;
RT "The nucleotide sequence of the mouse immunoglobulin epsilon gene:
RT comparison with the human epsilon gene sequence.";
RL EMBO J. 1:1117-1123(1982).
RN [2]
RP REVISIONS.
RA Honjo T.;
RL Submitted (APR-1986) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 34-421 FROM N.A.
RX MEDLINE=83117774; PubMed=6818553;
RA Liu F.-T., Albrandt K., Sutcliffe J.G., Katz D.H.;
RT "Cloning and nucleotide sequence of mouse immunoglobulin epsilon
RT chain cDNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:7852-7856(1982).
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 08:17:40 ; Search time 6.4 Seconds  
(without alignments)  
203.399 Million cell updates/sec

Title: US-09-701-623C-84

Perfect score: 135

Sequence: 1 CGETYKSTVSHPDLPREVRSIAKC 25

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	71	52.6	428	1	EPC_HUMAN
2	66	48.9	421	1	EPC_MOUSE
3	66	48.9	429	1	EPC_RAT
4	54	40.0	597	1	LEPA_VIBPA
5	54	40.0	597	1	LEPA_VIBVU
6	54	40.0	598	1	LEPA_HAEN
7	54	40.0	598	1	LEPA_PASMU
8	54	40.0	599	1	LEPA_ECOLI
9	54	40.0	599	1	LEPA_SALTY
10	54	40.0	602	1	LEPA_CHLMU
11	54	40.0	602	1	LEPA_CHLTR
12	52.5	38.9	450	1	MUC_CANPA
13	52	38.5	606	1	LEPA_CAUCR
14	52	38.5	607	1	LEPA_STAEP
15	51	37.8	340	1	ALC2_HUMAN
16	51	37.8	391	1	MUCB_HUMAN
17	51	37.8	454	1	MUC_HUMAN
18	51	37.8	457	1	MUC_SUNMU
19	51	37.8	458	1	MUC_RABIT
20	51	37.8	479	1	MUCM_RABIT
21	51	37.8	597	1	LEPA_VIBCH
22	51	37.8	599	1	LEPA_YERPE
23	50	37.0	599	1	LEPA_PSEAE
24	50	37.0	599	1	LEPA_PSEPK
25	50	37.0	600	1	LEPA_AQUAE
26	50	37.0	603	1	LEPA_SYNEL
27	50	37.0	607	1	LEPA_LACLA
28	50	37.0	607	1	LEPA_STRPN
29	50	37.0	607	1	LEPA_STR6
30	50	37.0	608	1	LEPA_LISIN
31	50	37.0	608	1	LEPA_LISMO
32	50	37.0	610	1	LEPA_STR3
33	50	37.0	610	1	LEPA_STRP3
					P01854 homo sapien
					P06336 mus musculus
					P01855 rattus norv
					Q871n7 vibrio para
					Q8dc78 vibrio vuln
					P43729 haemophilus
					P57806 pasteurella
					P07882 escherichia
					P23598 salmonella
					Q9pkx6 chlamydia t
					Q84057 chlamydia t
					P01874 canis fami
					Q9a9f4 caulobacter
					Q8c9l3 staphylococ
					P01877 homo sapien
					P04220 homo sapien
					P01871 homo sapien
					P20768 suncus muri
					P03988 oryctolagus
					P04221 oryctolagus
					Q9kb04 vibrio chol
					Q8zd74 yersinia pe
					Q9c9j8 lactococcus
					Q9dk5 streptococ
					Q8dph5 streptococ
					Q92bn4 listeria m
					Q8y742 listeria m
					Q8e5r3 streptococ
					Q8k7m8 streptococ

#### ALIGNMENTS

##### RESULT 1

EPC_HUMAN	ID	EPC_HUMAN	STANDARD;	PRT;	428 AA.
AC	P01854				
DT	21-JUL-1986	(Rel. 01, Created)			
DT	21-JUL-1986	(Rel. 01, Last sequence update)			
DT	10-OCT-2003	(Rel. 42, Last annotation update)			
DE	Ig epsilon chain C region.				
GN	IGHE				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=83168897; PubMed=6300763;				
RA	Seno M., Kurokawa T., Ono Y., Onda H., Sasada R., Igarashi K.,				
RA	Kikuchi M., Sugino Y., Nishida Y., Honjo T.;				
RT	"Molecular cloning and nucleotide sequencing of human immunoglobulin epsilon chain cDNA.";				
RL	Nucleic Acids Res. 11:719-726 (1983).				
RN	[2]				
RP	SEQUENCE FROM N.A.; AND VARIANT LEU-359.				
RX	MEDLINE=83001945; PubMed=6288268;				
RA	Max E.E., Battey J.; Ney R., Kirsch I.R., Leder P.;				
RT	"Duplication and deletion in the human immunoglobulin epsilon genes.";				
RL	Cell 29:691-699 (1982).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=84236029; PubMed=6234164;				
RA	Flanagan J.G., Rabbitts T.H.;				
RT	"The sequence of a human immunoglobulin epsilon heavy chain constant region gene, and evidence for three non-allelic genes.";				
RL	EMBO J. 1:655-660 (1982).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=84207910; PubMed=6327276;				
RA	Ueda S., Nakai S., Nishida Y., Hisajima H., Honjo T.;				
RT	"Long terminal repeat-like elements flank a human immunoglobulin epsilon pseudogene that lacks introns.";				
RL	EMBO J. 1:1539-1544 (1982).				
RN	[5]				
RP	PRELIMINARY SEQUENCE (MYELOMA PROTEIN ND).				
RA	Bennich H.H., Johanson S.G.O., von Bahr-Lindstrom H.;				
RL	(In) Bach M.K. (eds.);				
RL	Immediate hypersensitivity: modern concepts and developments, pp.1-36,				
RL	Marcel Dekker, New York (1978).				
RN	[6]				
RP	SEQUENCE OF 1-40; 68-114 AND 427-428 FROM N.A.				
RX	MEDLINE=83065234; PubMed=6815656;				
RA	Kenten J.H., Molgaard H.V., Houghton M., Derbyshire R.B., Viney J.,				
RA	Bell L.O., Gould H.J.;				
RT	"Cloning and sequence determination of the gene for the human immunoglobulin epsilon chain expressed in a myeloma cell line.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 79:6661-6665 (1982).				
RN	[7]				

Q92zv8 streptococ  
Q8evn0 lactobacill  
Q8dtf3 streptococ  
Q02100 saccharomyc  
P01863 mus musculu  
P01866 mus musculu  
Q828a4 salmonella  
Q828a4 salmonella  
Q82qr6 salmonella  
P01865 mus musculu  
P01867 mus musculu  
Q9pmv3 xanthomonas  
Q823h7 chlamydomphi

34 50 37.0 610 1 LEPA\_STRPY  
35 50 37.0 611 1 LEPA\_LACPL  
36 50 37.0 611 1 LEPA\_STRMU  
37 50 37.0 647 1 SKO1\_YEAST  
38 49 36.3 330 1 GCAA\_MOUSE  
39 49 36.3 336 1 GCB\_MOUSE  
40 49 36.3 352 1 MODC\_SALTY  
41 49 36.3 352 1 MODC\_SALTY  
42 49 36.3 399 1 GCAM\_MOUSE  
43 49 36.3 405 1 GCAM\_MOUSE  
44 49 36.3 601 1 LEPA\_XANAC  
45 49 36.3 602 1 LEPA\_CHLCV

```
A:Molecule type: DNA
A:Residues: 1-599 <BLAT>
A:Cross-references: GB:AE000343; GB:U00096; NID:G2367139; PIDN:AAC75622.1; PID:gl1788922;
A:Accession: A85903
A:Experimental source: strain K-12, substrain MG1655
R:March, P.E.; Inouye, M.
J. Biol. Chem. 260, 7206-7213, 1985
A:Title: Characterization of the lep operon of Escherichia coli. Identification of the P
A:Reference number: A22627; MUID:85207751; PMID:2987248
A:Accession: A22627
A:Molecule type: DNA
A:Residues: 1-52,'N',54-155,'HRRGALFSENRWCAGRSR',156,'SGA',160,'H',162-163,'AGS',186-26
A:Note: the authors translated the codon AAC for residue 53 as Thr
C:Comment: The overproduction of this protein is lethal to E. coli. It is present in the
codes for signal peptidase.
C:Genetics:
A:Gene: lepA
A:Map position: 55 min
C:Superfamily: GTP-binding membrane protein lepA; translation elongation factor Tu homol
C:Keywords: GTP binding; membrane protein; nucleotide binding; P-loop
P:5-134/Domain: translation elongation factor Tu homology <ETU>
P:11-18/Region: nucleotide-binding motif A (P-loop)
P:131-134/Region: GTP-binding NKX motif
P:162-164/Region: GTP-binding SAK/L motif
P:17,18,53,131,132,134,162/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #sta
Query Match 40.0%; Score 54; DB 1; Length 599;
Best Local Similarity 46.4%; Pred. No. 5.2;
Matches 13; Conservative 2; Mismatches 9; Indels 4; Gaps 1;
Qy 2 GETYK----STVSHPDLPREVVRSLAK 25
Db 68 GETYQLNFIDTPGHVDFSEVSRSLAAC 95
RESULT 14
C91058
GTP-binding elongation factor [imported] - Escherichia coli (strain O157:H7, substrain R
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: C91058
R:Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gend
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: C91058
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-599 <HAY>
A:Cross-references: PIDN:BA836858.1; PID:gl13362906; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: EC83435
C:Superfamily: GTP-binding membrane protein lepA; translation elongation factor Tu homol
Query Match 40.0%; Score 54; DB 2; Length 599;
Best Local Similarity 46.4%; Pred. No. 5.2;
Matches 13; Conservative 2; Mismatches 9; Indels 4; Gaps 1;
Qy 2 GETYK----STVSHPDLPREVVRSLAK 25
Db 68 GETYQLNFIDTPGHVDFSEVSRSLAAC 95
RESULT 15
A85903
GTP-binding membrane protein lepA - Escherichia coli (strain O157:H7, substrain BDL933)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 27-Nov-2001
C:Accession: A85903
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Hiller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: A85903
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-599 <STO>
A:Cross-references: GB:AE005174; NID:gl2516987; PIDN:AAG57685.1; GSPDB:GN00145; UWGP:Z31
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: lepA
C:Superfamily: GTP-binding membrane protein lepA; translation elongation factor Tu homol
Query Match 40.0%; Score 54; DB 2; Length 599;
Best Local Similarity 46.4%; Pred. No. 5.2;
Matches 13; Conservative 2; Mismatches 9; Indels 4; Gaps 1;
Qy 2 GETYK----STVSHPDLPREVVRSLAK 25
Db 68 GETYQLNFIDTPGHVDFSEVSRSLAAC 95
Search completed: February 26, 2004, 08:28:17
Job time : 11 secs
```

Ig Y heavy chain (7.8S) - duck  
 N:Alternate names: Ig gamma chain (7.8S)  
 C:Species: Anas platyrhynchos (domestic duck)  
 C>Date: 18-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 21-Jul-2000  
 C:Accession: B46529; S20759  
 R:Magor, K.B.; Warr, G.W.; Middleton, D.; Wilson, M.R.; Higgins, D.A.  
 J:Immunol. 149, 2627-2633, 1992  
 A:Title: Structural relationship between the two IgY of the duck, Anas platyrhynchos: molecular cloning and characterization of the heavy chain cDNA  
 A:Reference number: A46529; MUID:93017865; PMID:1401901  
 A:Accession: B46529  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-572 <AG>  
 A:Cross-references: EMBL:X65219; NID:962442; PIDN:CAA46322.1; PID:962443  
 A:Experimental source: spleen  
 A>Note: sequence extracted from NCBI backbone (NCBIP:116127)  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: immunoglobulin  
 F:37-120/Domain: immunoglobulin homology <IMM>  
 Query Match 42.2%; Score 57; DB 2; Length 572;  
 Best Local Similarity 52.2%; Pred. No. 1.8;  
 Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;  
 QY 2 GETYKSTVSHDPLPREVRSIAK 24  
 DB 437 GERTCTVQHEDLPVPLGKSIK 459  
 RESULT 10  
 S70214  
 lepA protein homology - Photobacterium sp. (fragment)  
 C:Species: Photobacterium sp.  
 C>Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 02-Feb-2001  
 C:Accession: S70214  
 R:Chi, E.; Bartlett, D.H.  
 Mol. Microbiol. 17, 713-726, 1995  
 A:Title: An rpoS-like locus controls outer membrane protein synthesis and growth at cold temperatures in Photobacterium  
 A:Reference number: S70210; MUID:96111491; PMID:8801425  
 A:Accession: S70214  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-137 <CHI>  
 A:Cross-references: EMBL:L41667; NID:977746; PIDN:AAB31927.1; PID:9777751  
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1995  
 C:Superfamily: GTP-binding membrane protein lepA; translation elongation factor Tu homology  
 C:Keywords: GTP binding; nucleotide binding; P-loop  
 F:5-134/Domain: translation elongation factor Tu homology <ETU>  
 F:11-18/Region: nucleotide-binding motif A (P-loop)  
 F:11-18/Region: translation elongation factor Tu homology  
 F:131-134/Region: GTP-binding NKKD motif  
 Query Match 40.0%; Score 54; DB 2; Length 137;  
 Best Local Similarity 46.4%; Pred. No. 1.1;  
 Matches 13; Conservative 2; Mismatches 9; Indels 4; Gaps 1;  
 QY 2 GETYK----STVSHDPLPREVRSIAK 25  
 DB 68 GETYQLNFIDTPGHVDFSYEVSRLAAC 95  
 RESULT 11  
 EHMSS  
 Ig epsilon chain C region (version 2) - mouse  
 C:Species: Mus musculus (house mouse)  
 C>Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 16-Aug-1996  
 C:Accession: A02145  
 R:Ishida, N.; Ueda, S.; Hayashida, H.; Miyata, T.; Honjo, T.  
 EMBO J. 1, 1117-1123, 1982  
 A:Title: The nucleotide sequence of the mouse immunoglobulin epsilon chain with reference to the structure of the heavy chain constant region  
 A:Reference number: A09966; MUID:84236092; PMID:6329728  
 A:Accession: A02145  
 A:Molecule type: DNA  
 A:Residues: 1-423 <ISH>

A>Note: the sequence was determined from the germline gene

C:Genetics:  
 A:Introns: 91/1; 199/1; 307/1  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into larger oligomers.  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
 F:16-77/Domain: immunoglobulin homology <IMM1>  
 F:115-183/Domain: immunoglobulin homology <IMM2>  
 F:220-288/Domain: immunoglobulin homology <IMM3>  
 F:235-396/Domain: immunoglobulin homology <IMM4>  
 F:323-75,122-181,227-286,332-394/Disulfide bonds: #status predicted  
 F:43,84,167,239,262,417/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 Query Match 40.0%; Score 54; DB 1; Length 423;  
 Best Local Similarity 47.6%; Pred. No. 3.6;  
 Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;  
 QY 2 GETYKSTVSHDPLPREVRSI 22  
 DB 281 GYGQCVDVDRDPFKPIVRSI 301  
 RESULT 12  
 I64042  
 GTP-binding membrane protein lepA - Haemophilus influenzae (strain Rd KW20)  
 C:Species: Haemophilus influenzae  
 C>Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 02-Feb-2001  
 C:Accession: I64042  
 R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, J.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.F.; Glodek, A.; Kelley, J.M.; Weidman, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.; Science 269, 496-512, 1995  
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
 A:Reference number: A64000; MUID:95350830; PMID:7542800  
 A:Accession: I64042  
 A>Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-598 <TIGR>  
 A:Cross-references: GB:U32687; GB:L42023; NID:91572955; PIDN:AAC21694.1; PID:91572960; C:Genetics:  
 A:Gene: lepA  
 C:Superfamily: GTP-binding membrane protein lepA; translation elongation factor Tu homology  
 C:Keywords: GTP binding; membrane protein; nucleotide binding; P-loop  
 F:5-134/Domain: translation elongation factor Tu homology <ETU>  
 F:11-18/Region: nucleotide-binding motif A (P-loop)  
 F:131-134/Region: GTP-binding NKKD motif  
 F:162-164/Region: GTP-binding SAK/L motif  
 F:17,18,53,131,132,134,162/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #st:  
 Query Match 40.0%; Score 54; DB 2; Length 598;  
 Best Local Similarity 46.4%; Pred. No. 5.2;  
 Matches 13; Conservative 2; Mismatches 9; Indels 4; Gaps 1;  
 QY 2 GETYK----STVSHDPLPREVRSIAK 25  
 DB 68 GETYQLNFIDTPGHVDFSYEVSRLAAC 95  
 RESULT 13  
 BVECLA  
 GTP-binding membrane protein lepA - Escherichia coli (strain K-12)  
 C:Species: Escherichia coli  
 C>Date: 31-Mar-1988 #sequence\_revision 05-Dec-1997 #text\_change 01-Mar-2002  
 C:Accession: H65034; A2627  
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C.A.; Rose, D.J.; Mau, B.; Shao, Y.  
 Science 277, 1453-1462, 1997  
 A:Title: The complete genome sequence of Escherichia coli K-12.  
 A:Reference number: A64720; MUID:97426617; PMID:9278503  
 A:Accession: H65034  
 A>Status: nucleic acid sequence not shown; translation not shown



A:Accession: I68726  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-107 <RSS>  
 A:Cross-references: GB:W22930; NID:G194455; PIDN:AAA37911.1; PID:G194460  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 F:22-90/Domain: immunoglobulin homology <IMM>

Query Match 48.9%; Score 66; DB 2; Length 107;  
 Best Local Similarity 52.2%; Pred. No. 0.014;  
 Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 GETYKSTVSHDPLPREVVRSTAK 24  
 |||:|||||:|||||  
 Db 83 GYGQCIVDHPDFPKPIVRSITK 105

## RESULT 5

EHMS

Ig epsilon chain C region (version 1) - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 13-Jun-1983 #sequence\_revision 13-Jun-1983 #text\_change 16-Jul-1999  
 C:Accession: A02144  
 R:Liu, F.T.; Albrandt, K.; Sutcliffe, J.G.; Katz, D.H.  
 Proc. Natl. Acad. Sci. U.S.A. 79, 7852-7856, 1982  
 A:Title: Cloning and nucleotide sequence of mouse immunoglobulin epsilon chain cDNA.  
 A:Reference number: A02144; MUID:83117774; PMID:6818553  
 A:Accession: A02144  
 A:Molecule type: mRNA  
 A:Residues: 1-388 <LIU>  
 A:Cross-references: GB:J00476; NID:G194875; PIDN:AAA38085.1; PID:G387220  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap) chain disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into la  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
 F:1-44/Domain: immunoglobulin homology (fragment) <IM1>  
 F:81-149/Domain: immunoglobulin homology <IM2>  
 F:186-254/Domain: immunoglobulin homology <IM3>  
 F:290-361/Domain: immunoglobulin homology <IM4>  
 F:10,51,62,133,205,228,332,382/Binding site: carbohydrate (Asn) (covalent) #status predict

Query Match 48.9%; Score 66; DB 1; Length 388;  
 Best Local Similarity 52.2%; Pred. No. 0.053;  
 Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 GETYKSTVSHDPLPREVVRSTAK 24  
 |||:|||||:|||||  
 Db 247 GYGQCIVDHPDFPKPIVRSITK 269

## RESULT 6

EHRT

Ig epsilon chain C region - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 17-Dec-1982 #sequence\_revision 17-Dec-1982 #text\_change 16-Jul-1999  
 C:Accession: A93442; A90937; A02143  
 R:Hellman, L.; Petersson, U.; Engstrom, A.; Karlsson, T.; Bennich, H.  
 Nucleic Acids Res. 10, 6041-6049, 1982  
 A:Title: Structure and evolution of the heavy chain from rat immunoglobulin E.  
 A:Reference number: A93442; MUID:83064537; PMID:6292865  
 A:Accession: A93442  
 A:Molecule type: mRNA  
 A:Residues: 1-429 <HEL>  
 A:Experimental source: strain LOU/c/Wsl, immunocytoma IR2  
 R:Kindevoegel, W.R.; Reddy, E.P.; Moore, J.M.; Faust Jr., C.H.  
 DNA 1, 335-343, 1982  
 A:Title: A cloned cDNA probe for rat immunoglobulin epsilon heavy chain: construction, i  
 A:Reference number: A90937; MUID:83182019; PMID:6820340  
 A:Contents: myeloma IR162  
 A:Accession: A90937  
 A:Molecule type: mRNA  
 A:Residues: 'N',189-307,'L',309-342 <KIN>  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap

chain disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into la  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
 F:19-80/Domain: immunoglobulin homology <IM1>  
 F:118-186/Domain: immunoglobulin homology <IM2>  
 F:223-291/Domain: immunoglobulin homology <IM3>  
 F:327-398/Domain: immunoglobulin homology <IM4>  
 F:46,99,170,240,265,369,419/Binding site: carbohydrate (Asn) (covalent) #status predict;

Query Match 48.9%; Score 66; DB 1; Length 429;  
 Best Local Similarity 52.2%; Pred. No. 0.059;  
 Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 GETYKSTVSHDPLPREVVRSTAK 24  
 |||:|||||:|||||  
 Db 284 GEGYQCRVDHPDFPKPIVRSITK 306

## RESULT 7

S38864  
 Ig epsilon chain C region - mouse (fragment)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 24-May-2001  
 C:Accession: S38864  
 R:Kipp, B.; Becker, W.; Schlaak, M.  
 submitted to the EMBL Data Library, November 1993  
 A:Description: Combination of a defined specificity and desired isotype by cloning of a  
 A:Reference number: S38864  
 A:Accession: S38864  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-548 <KIP>  
 A:Cross-references: EMBL:Z27397; NID:G416537; PIDN:CAA81788.1; PID:G940782  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 F:353-421/Domain: immunoglobulin homology <IMW>

Query Match 48.9%; Score 66; DB 2; Length 548;  
 Best Local Similarity 52.2%; Pred. No. 0.076;  
 Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 GETYKSTVSHDPLPREVVRSTAK 24  
 |||:|||||:|||||  
 Db 414 GYGQCIVDHPDFPKPIVRSITK 436

## RESULT 8

S43147

Ig upsilon chain - duck (fragment)  
 C:Species: Anas platyrhynchos (domestic duck)  
 C:Date: 06-Feb-1995 #sequence\_revision 06-Feb-1995 #text\_change 11-Jan-2000  
 C:Accession: S43147  
 R:Nagor, K.E.; Higgins, D.A.; Middleton, D.L.; Warr, G.W.  
 submitted to the EMBL Data Library, March 1994  
 A:Description: Evidence from duck immunoglobulin genes that Igy is the common ancestor  
 A:Reference number: S43145  
 A:Accession: S43147  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-110 <MAG>  
 A:Cross-references: EMBL:X78355; NID:G468612; PID:G468613  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 44.4%; Score 60; DB 2; Length 110;  
 Best Local Similarity 52.2%; Pred. No. 0.11;  
 Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 GETYKSTVSHDPLPREVVRSTAK 24  
 |||:|||||:|||||  
 Db 86 GERFTCTVQHEDLPEPLGKSIK 108

## RESULT 9

B46529



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 08:18:56 ; Search time 10 Seconds  
(without alignments)  
240.479 Million cell updates/sec

Title: US-09-701-623C-84

Perfect score: 135

Sequence: 1 CGETYKSTVSHPDLPREVRSIAKC 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:\*  
1: PIR1.\*  
2: PIR2.\*  
3: PIR3.\*  
4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	74	54.8	426	2	Ig epsilon-chain -
2	71	52.6	428	1	Ig epsilon chain C
3	66	48.9	107	2	Ig chain C3 regio
4	66	48.9	107	2	Ig chain C3 regio
5	66	48.9	388	1	Ig epsilon chain C
6	66	48.9	429	1	Ig epsilon chain C
7	66	48.9	548	2	Ig epsilon chain C
8	60	44.4	110	2	Ig epsilon chain C
9	57	42.2	572	2	Ig epsilon chain C
10	54	40.0	137	2	Ig epsilon chain C
11	54	40.0	423	1	Ig epsilon chain C
12	54	40.0	598	2	Ig epsilon chain C
13	54	40.0	599	1	Ig epsilon chain C
14	54	40.0	599	2	Ig epsilon chain C
15	54	40.0	599	2	Ig epsilon chain C
16	54	40.0	599	2	Ig epsilon chain C
17	54	40.0	602	2	Ig epsilon chain C
18	54	40.0	602	2	Ig epsilon chain C
19	53	39.3	684	2	Ig epsilon chain C
20	52.5	38.9	450	1	Ig epsilon chain C
21	52	38.5	394	2	Ig epsilon chain C
22	52	38.5	648	2	Ig epsilon chain C
23	51	37.8	93	2	Ig epsilon chain C
24	51	37.8	220	2	Ig epsilon chain C
25	51	37.8	340	1	Ig epsilon chain C
26	51	37.8	340	2	Ig epsilon chain C
27	51	37.8	340	2	Ig epsilon chain C
28	51	37.8	391	1	Ig epsilon chain C
29	51	37.8	452	1	Ig epsilon chain C

Ig mu chain C regi  
Ig mu chain C regi  
Ig mu chain C regi  
Ig mu chain C regi  
Ig heavy chain - h  
Ig heavy chain - h  
Ig mu chain C regi  
GTP-binding protei  
probable GTP-bind  
Ig mu chain precu  
hypothetical prote  
Ig gamma 3 chain c  
Ig gamma 1 chain c  
GTP-binding protei  
G-protein lepa - A  
GTP-binding protei  
GTP-binding protei

#### ALIGNMENTS

##### RESULT 1

I36948  
Ig epsilon-chain - chimpanzee (fragment)  
C:Species: Pan troglodytes (chimpanzee)  
C:Date: 04-Oct-1995 #sequence\_revision 04-Oct-1996 #text\_change 21-Jan-2000  
C:Accession: I36948  
R:Sakoyama, Y.; Hong, K.  
Proc. Natl. Acad. Sci. U.S.A. 84, 1080-1084, 1987  
A:Title: Nucleotide sequences of immunoglobulin epsilon genes of chimpanzee and orangutan  
A:Reference number: I36948; MUID:87147196; PMID:3103123  
A:Accession: I36948  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-426 <RES>  
A:Cross-references: GB:M15398; NID:gl76797; PIDN:AAA35416.1; PID:gl76798  
C:Genetics:  
A:Introns: 103/1; 209/1; 317/1  
C:Superfamily: immunoglobulin C region; immunoglobulin homology  
F:336-405/Domain: immunoglobulin homology <IMM>

Query Match 54.8%; Score 74; DB 2; Length 426;  
Best Local Similarity 60.9%; Pred. No. 0.0037; 6; Indels 0; Gaps 0;  
Matches 14; Conservative 3; Mismatches 6

QY 2 GETYKSTVSHPDLPREVRSIAK 24  
|||||:|||||:|||||  
DB 292 GETYQCVTHPHLPALVRSTTK 314

##### RESULT 2

EHHU  
Ig epsilon chain C region - human  
C:Species: Homo sapiens (man)  
C:Date: 31-Mar-1981 #sequence\_revision 13-Jun-1983 #text\_change 16-Jul-1999  
C:Accession: A22771; A23195; PH1214; A93491; A90824; A94418; B93933; S02438; A53116; C4  
R:Flanagan, J.G.; Rabbitts, T.H.  
EMBO J. 1, 655-660, 1982  
A:Title: The sequence of a human immunoglobulin epsilon heavy chain constant region gen  
A:Reference number: A22771; MUID:84236029; PMID:6234164  
A:Accession: A22771  
A:Molecule type: DNA  
A:Residues: 1-428 <FLA>  
A:Cross-references: GB:I00022; GB:J00227; GB:V00555; NID:gl85035  
R:Ueda, S.; Nakai, S.; Nishida, Y.; Hisejima, H.; Honjo, T.  
EMBO J. 1, 1539-1544, 1982  
A:Title: Long terminal repeat-like elements flank a human immunoglobulin epsilon pseudo  
A:Reference number: A23195; MUID:84207910; PMID:6327276  
A:Accession: A23195  
A:Molecule type: DNA  
A:Residues: 2-428 <UED>  
A:Cross-references: GB:J00222; NID:gl84755  
R:Zhang, K.; Saxon, A.; Max, E.E.

ADDRESSEE: Morgan & Finnegan, L.L.P.  
STREET: 345 Park Avenue  
CITY: New York

```
RESULT 11
US-09-770-014-98
; Sequence 98, Application US/09770014
; Patent No. 6559282
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/770,014
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; NAME: Maria H. Lin
; ATTORNEY/AGENT INFORMATION:
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-770-014-98

Query Match 45.7%; Score 64; DB 4; Length 42;
Best Local Similarity 48.0%; Pred. No. 0.0064;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYGQSVDRDPFKPIVRSTLC 25
Db 18 CGETQSRVTHPLPALMRSTTKC 42

RESULT 12
US-09-770-014-98
; Sequence 99, Application US/09770014
; Patent No. 6559282
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/770,014
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; NAME: Maria H. Lin
; ATTORNEY/AGENT INFORMATION:
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 99:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-770-014-99

Query Match 45.7%; Score 64; DB 4; Length 42;
Best Local Similarity 48.0%; Pred. No. 0.0064;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYGQSVDRDPFKPIVRSTLC 25
Db 18 CGETQSRVTHPLPALMRSTTKC 42

RESULT 13
US-09-770-014-100
; Sequence 100, Application US/09770014
; Patent No. 6559282
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/770,014
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; NAME: Maria H. Lin
; ATTORNEY/AGENT INFORMATION:
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 100:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
```

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Morgan & Finnegan, L.L.P.  
;; STREET: 345 Park Avenue  
;; CITY: New York  
;; STATE: NY  
;; COUNTRY: USA  
;; ZIP: 10154-0054  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC Windows  
;; SOFTWARE: Word 97  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/303,323  
;; FILING DATE: 30-APR-1999  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 09/100,414  
;; FILING DATE: 20-JUNE-1998  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Maria H. Lin  
;; REGISTRATION NUMBER: 29,323  
;; REFERENCE/DOCKET NUMBER: 1151-4157  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 212-758-4800  
;; TELEFAX: 212-751-6849  
;; INFORMATION FOR SEQ ID NO: 98:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 42 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-09-303-323-98

Query Match 45.7%; Score 64; DB 3; Length 42;  
Best Local Similarity 48.0%; Pred. No. 0.0064;  
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYGQSIYDRDPFKPIVRSITLC 25  
DB 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 9  
US-09-303-323-99  
; Sequence 99, Application US/09303323  
; Patent No. 6228987  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Chang Yi  
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE  
; TITLE OF INVENTION: IMMUNOGENS  
; NUMBER OF SEQUENCES: 106  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Morgan & Finnegan, L.L.P.  
; STREET: 345 Park Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10154-0054  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC Windows  
; SOFTWARE: Word 97  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/303,323  
; FILING DATE: 30-APR-1999  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/100,414  
; FILING DATE: 20-JUNE-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maria H. Lin  
; REGISTRATION NUMBER: 29,323  
; REFERENCE/DOCKET NUMBER: 1151-4157  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-758-4800  
; TELEFAX: 212-751-6849  
; INFORMATION FOR SEQ ID NO: 98:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 42 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide

;; REGISTRATION NUMBER: 29,323  
;; REFERENCE/DOCKET NUMBER: 1151-4157  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 212-758-4800  
;; TELEFAX: 212-751-6849  
;; INFORMATION FOR SEQ ID NO: 99:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 42 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-09-303-323-99

Query Match 45.7%; Score 64; DB 3; Length 42;  
Best Local Similarity 48.0%; Pred. No. 0.0064;  
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYGQSIYDRDPFKPIVRSITLC 25  
DB 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 10  
US-09-303-323-100  
; Sequence 100, Application US/09303323  
; Patent No. 6228987  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Chang Yi  
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE  
; TITLE OF INVENTION: IMMUNOGENS  
; NUMBER OF SEQUENCES: 106  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Morgan & Finnegan, L.L.P.  
; STREET: 345 Park Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10154-0054  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC Windows  
; SOFTWARE: Word 97  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/303,323  
; FILING DATE: 30-APR-1999  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/100,414  
; FILING DATE: 20-JUNE-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maria H. Lin  
; REGISTRATION NUMBER: 29,323  
; REFERENCE/DOCKET NUMBER: 1151-4157  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-758-4800  
; TELEFAX: 212-751-6849  
; INFORMATION FOR SEQ ID NO: 100:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 42 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-303-323-100

Query Match 45.7%; Score 64; DB 3; Length 42;  
Best Local Similarity 48.0%; Pred. No. 0.0064;  
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYGQSIYDRDPFKPIVRSITLC 25  
DB 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 8  
US-09-303-323-98  
; Sequence 98, Application US/09303323  
; Patent No. 6228987  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Chang Yi  
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE  
; TITLE OF INVENTION: IMMUNOGENS  
; NUMBER OF SEQUENCES: 106

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-95

Query Match 45.7%; Score 64; DB 3; Length 25;
Best Local Similarity 48.0%; Pred. No. 0.0036;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYGQSVIVDRDPFKPIVRSITLC 25
Db 1 CGETQSVRTHPLPALMRSTTKC 25

RESULT 3
US-09-303-323-95
; Sequence 95, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

Query Match 45.7%; Score 64; DB 3; Length 25;
Best Local Similarity 48.0%; Pred. No. 0.0036;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYGQSVIVDRDPFKPIVRSITLC 25
Db 1 CGETQSVRTHPLPALMRSTTKC 25

US-09-303-323-95
; Sequence 95, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

```

```

US-09-303-323-95
Query Match 45.7%; Score 64; DB 3; Length 25;
Best Local Similarity 48.0%; Pred. No. 0.0036;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYGQSVIVDRDPFKPIVRSITLC 25
Db 1 CGETQSVRTHPLPALMRSTTKC 25

RESULT 4
US-09-770-014-95
; Sequence 95, Application US/09770014
; Patent No. 6559282
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/770,014
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-770-014-95

Query Match 45.7%; Score 64; DB 4; Length 25;
Best Local Similarity 48.0%; Pred. No. 0.0036;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYGQSVIVDRDPFKPIVRSITLC 25
Db 1 CGETQSVRTHPLPALMRSTTKC 25

RESULT 5
US-09-100-414B-98
; Sequence 98, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 08:19:31 ; Search time 12.6 Seconds  
(without alignments)  
102.432 Million cell updates/sec

Title: US-09-701-623C-8  
Perfect score: 140  
Sequence: 1 GYGQSVDRPDPFKPIVRSITL 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*

- 1: /cgm2\_6/ptodata/2/iaa/5A\_COMB.pap:\*
- 2: /cgm2\_6/ptodata/2/iaa/5B\_COMB.pap:\*
- 3: /cgm2\_6/ptodata/2/iaa/6A\_COMB.pap:\*
- 4: /cgm2\_6/ptodata/2/iaa/6B\_COMB.pap:\*
- 5: /cgm2\_6/ptodata/2/iaa/PTUS\_COMB.pap:\*
- 6: /cgm2\_6/ptodata/2/iaa/backfiles.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	116	82.9	561	3	US-09-192-545-2
2	64	45.7	25	3	US-09-100-414B-95
3	64	45.7	25	3	US-09-303-323-95
4	64	45.7	25	4	US-09-770-014-95
5	64	45.7	42	3	US-09-100-414B-98
6	64	45.7	42	3	US-09-100-414B-99
7	64	45.7	42	3	US-09-100-414B-100
8	64	45.7	42	3	US-09-303-323-98
9	64	45.7	42	3	US-09-303-323-99
10	64	45.7	42	3	US-09-303-323-100
11	64	45.7	42	4	US-09-770-014-98
12	64	45.7	42	4	US-09-770-014-99
13	64	45.7	42	4	US-09-770-014-100
14	64	45.7	45	3	US-09-100-414B-101
15	64	45.7	45	3	US-09-303-323-101
16	64	45.7	45	4	US-09-770-014-101
17	64	45.7	46	3	US-09-100-414B-96
18	64	45.7	46	3	US-09-100-414B-97
19	64	45.7	46	3	US-09-303-323-96
20	64	45.7	46	3	US-09-303-323-97
21	64	45.7	46	4	US-09-770-014-96
22	64	45.7	46	4	US-09-770-014-97
23	64	45.7	63	3	US-09-100-414B-102
24	64	45.7	63	3	US-09-303-323-102
25	64	45.7	63	4	US-09-770-014-102
26	58	41.4	431	4	US-09-479-614-14
27	58	41.4	496	4	US-09-479-614-2

28 41.4 496 4 US-09-479-614-29 Sequence 29, Appl  
48 34.3 226 4 US-09-252-991A-25458 Sequence 25458, A  
30 48 34.3 268 3 US-09-187-049-9 Sequence 9, Appl  
31 48 34.3 1259 3 US-09-187-049-13 Sequence 13, Appl  
32 47 33.6 202 4 US-09-252-991A-17334 Sequence 17334, A  
33 46 32.9 150 4 US-09-252-991A-20542 Sequence 20542, A  
34 46 32.9 282 4 US-09-252-991A-21097 Sequence 21097, A  
35 46 32.9 919 4 US-09-543-681A-4258 Sequence 4258, Ap  
36 45 32.1 426 5 US-08-336-583-2 Sequence 2, Appl  
37 45 32.1 426 5 PCT-US95-13795-2 Sequence 2, Appl  
38 45 32.1 544 4 US-09-543-681A-7663 Sequence 7663, Ap  
39 45 32.1 699 5 PCT-US94-07297-39 Sequence 39, Appl  
40 45 32.1 917 4 US-09-489-039A-12582 Sequence 12582, A  
41 45 32.1 921 1 US-08-396-479B-2 Sequence 2, Appl  
42 45 32.1 921 1 US-08-818-823-2 Sequence 2, Appl  
43 45 32.1 1046 4 US-09-540-236-2376 Sequence 2376, Ap  
44 45 32.1 1171 4 US-09-417-197-131 Sequence 131, App  
45 45 32.1 1181 4 US-09-417-197-133 Sequence 133, App

#### ALIGNMENTS

##### RESULT 1

US-09-192-545-2  
; Sequence 2, Application US/09192545  
; Patent No. 6118044  
; GENERAL INFORMATION:  
; APPLICANT: Karasuyama, Hajime  
; APPLICANT: Yonekawa, Hiromichi  
; APPLICANT: Taya, Choji  
; APPLICANT: Matsuoka, Kunie  
; TITLE OF INVENTION: Transgenic Animal Allergy Models and Methods for Their Use  
; FILE REFERENCE: 799P79570  
; CURRENT APPLICATION NUMBER: US/09/192,545  
; EARLIER FILING DATE: 1998-11-13  
; EARLIER APPLICATION NUMBER: JP HEI 9-313989  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 561  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE: Description of Artificial Sequence: Designed heavy

US-09-192-545-2

Query Match 82.9%; Score 116; DB 3; Length 561;  
Best Local Similarity 91.3%; Pred.No. 1.3e-09;  
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GYGQSVDRPDPFKPIVRSITL 24

Db 419 GYGQCVDRPDPFKPIVRSITL 441

##### RESULT 2

US-09-100-414B-95  
; Sequence 95, Application US/09100414B  
; Patent No. 6025468  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Chang Yi  
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE  
; TITLE OF INVENTION: IMMUNOGENS  
; NUMBER OF SEQUENCES: 106  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Morgan & Finnegan, L.L.P.  
; STREET: 345 Park Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10154-0054  
; COMPUTER READABLE FORM:

; OTHER INFORMATION: Murine IgE heavy chain, domains C1, C2, C3, C4, and MIGIS fragment  
; OTHER INFORMATION: t  
US-09-949-375A-19

Query Match 77.1%; Score 108; DB 9; Length 432;  
Best Local Similarity 90.9%; Pred. No. 4e-08;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GYGQSIIVDRDPDFPKPIVRSIT 23  
Db 280 GYGQCIIVDHPDFPKPIVRSIT 301

RESULT 12  
US-10-096-840D-2  
; Sequence 2, Application US/10096840D  
; Publication No. US20030158390A1  
; GENERAL INFORMATION:  
; APPLICANT: YISSUM Research Development Company of the Hebrew University of Jerusalem  
; TITLE OF INVENTION: Fce-PE CHIMERIC PROTEIN FOR TARGETED TREATMENT OF ALLERGY RESPONSES  
; FILE REFERENCE: 1268-067A  
; CURRENT APPLICATION NUMBER: US/10/096,840D  
; PRIOR FILING DATE: 2003-03-07  
; PRIOR FILING DATE: 1995-12-18  
; PRIOR APPLICATION NUMBER: US 09/091,645  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: PCT / IL96 / 00181  
; PRIOR FILING DATE: 1996-12-18  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 503  
; TYPE: PRT  
; ORGANISM: Mouse  
US-10-096-840D-2

Query Match 77.1%; Score 108; DB 14; Length 503;  
Best Local Similarity 90.9%; Pred. No. 4.7e-08;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GYGQSIIVDRDPDFPKPIVRSIT 23  
Db 113 GYGQCIIVDHPDFPKPIVRSIT 134

RESULT 13  
US-10-096-840D-4  
; Sequence 4, Application US/10096840D  
; Publication No. US20030158390A1  
; GENERAL INFORMATION:  
; APPLICANT: YISSUM Research Development Company of the Hebrew University of Jerusalem  
; TITLE OF INVENTION: Fce-PE CHIMERIC PROTEIN FOR TARGETED TREATMENT OF ALLERGY RESPONSES  
; FILE REFERENCE: 1268-067A  
; CURRENT APPLICATION NUMBER: US/10/096,840D  
; PRIOR FILING DATE: 2003-03-07  
; PRIOR APPLICATION NUMBER: IL 116436  
; PRIOR FILING DATE: 1995-12-18  
; PRIOR APPLICATION NUMBER: US 09/091,645  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: PCT / IL96 / 00181  
; PRIOR FILING DATE: 1996-12-18  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 4  
; LENGTH: 676  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-10-096-840D-4

Query Match 77.1%; Score 108; DB 14; Length 676;

Best Local Similarity 90.9%; Pred. No. 6.5e-08;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GYGQSIIVDRDPDFPKPIVRSIT 23  
Db 190 GYGQCIIVDHPDFPKPIVRSIT 211

RESULT 14  
US-10-152-190-5  
; Sequence 5, Application US/10152190  
; Publication No. US20030096369A1  
; GENERAL INFORMATION:  
; APPLICANT: Morsey, Mohamad A.  
; TITLE OF INVENTION: NO. US20030096369A1-anaphylactogenic IgE vaccines  
; FILE REFERENCE: PC11011A  
; CURRENT APPLICATION NUMBER: US/10/152,190  
; CURRENT FILING DATE: 2002-05-21  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 114  
; TYPE: PRT  
; ORGANISM: Rat CH3  
US-10-152-190-5

Query Match 60.7%; Score 85; DB 14; Length 114;  
Best Local Similarity 77.3%; Pred. No. 2.9e-05;  
Matches 17; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 GYGQSIIVDRDPDFPKPIVRSIT 23  
Db 84 GEGYQCRVDHPDFPKPIVRSIT 105

RESULT 15  
US-09-401-636-2  
; Sequence 2, Application US/09401636  
; Patent No. US20010038843A1  
; GENERAL INFORMATION:  
; APPLICANT: Hellman, Lars T.  
; TITLE OF INVENTION: ENHANCED VACCINES  
; FILE REFERENCE: 10223/006001  
; CURRENT APPLICATION NUMBER: US/09/401,636  
; CURRENT FILING DATE: 1999-09-22  
; PRIOR APPLICATION NUMBER: US 60/106,652  
; PRIOR FILING DATE: 1998-11-02  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 340  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetically generated proteins  
US-09-401-636-2

Query Match 60.7%; Score 85; DB 9; Length 340;  
Best Local Similarity 77.3%; Pred. No. 9.8e-05;  
Matches 17; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 GYGQSIIVDRDPDFPKPIVRSIT 23  
Db 195 GEGYQCRVDHPDFPKPIVRSIT 216

Search completed: February 26, 2004, 08:46:54  
Job time : 26.2 secs

```
US-09-949-375A-22
; Sequence 22, Application US/09949375A
; Patent No. US20020172673A1
; GENERAL INFORMATION:
; APPLICANT: KLYSNER, Steen et al.
; TITLE OF INVENTION: METHOD FOR DOWN-REGULATING IGE
; FILE REFERENCE: 3631-0111P
; CURRENT APPLICATION NUMBER: US/09/949,375A
; CURRENT FILING DATE: 2002-01-18
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial amino acid sequence of SEQ ID NO: 21.
US-09-949-375A-22
Query Match 77.1%; Score 108; DB 9; Length 343;
Best Local Similarity 90.9%; Pred. No. 3.1e-08;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 GYGQSIIVDRDPFKPIVRSIT 23
DB 191 GYGQCIIVDHPDFPKPIVRSIT 212
RESULT 8
US-09-949-375A-28
; Sequence 28, Application US/09949375A
; Patent No. US20020172673A1
; GENERAL INFORMATION:
; APPLICANT: KLYSNER, Steen et al.
; TITLE OF INVENTION: METHOD FOR DOWN-REGULATING IGE
; FILE REFERENCE: 3631-0111P
; CURRENT APPLICATION NUMBER: US/09/949,375A
; CURRENT FILING DATE: 2002-01-18
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 421
; TYPE: PRT
; ORGANISM: mus musculus
; FEATURE:
; NAME/KEY: MISC FEATURE
; OTHER INFORMATION: Murine Ige heavy chain domains C1, C2, C3, and C4.
US-09-949-375A-28
Query Match 77.1%; Score 108; DB 9; Length 421;
Best Local Similarity 90.9%; Pred. No. 3.9e-08;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 GYGQSIIVDRDPFKPIVRSIT 23
DB 280 GYGQCIIVDHPDFPKPIVRSIT 301
RESULT 9
US-10-214-524-31
; Sequence 31, Application US/10214524
; Publication No. US20030073142A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Sway-Shen Alex
; APPLICANT: Yang, Yong-Min
; APPLICANT: Barankiewicz, Theresa J.
; APPLICANT: Chen, Zhong
; TITLE OF INVENTION: IMMUNOGLOBULIN E VACCINES AND METHODS OF USE THEREOF
; FILE REFERENCE: IGE-00101.P.1.1
; CURRENT APPLICATION NUMBER: US/10/214,524
; CURRENT FILING DATE: 2002-08-08
; PRIOR FILING DATE: 2001-08-13
US-09-949-375A-19
; Sequence 19, Application US/09949375A
; Patent No. US20020172673A1
; GENERAL INFORMATION:
; APPLICANT: KLYSNER, Steen et al.
; TITLE OF INVENTION: METHOD FOR DOWN-REGULATING IGE
; FILE REFERENCE: 3631-0111P
; CURRENT APPLICATION NUMBER: US/09/949,375A
; CURRENT FILING DATE: 2002-01-18
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 432
; TYPE: PRT
; ORGANISM: mus musculus
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(432)
US-09-949-375A-19
Query Match 77.1%; Score 108; DB 14; Length 421;
Best Local Similarity 90.9%; Pred. No. 3.9e-08;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 GYGQSIIVDRDPFKPIVRSIT 23
DB 280 GYGQCIIVDHPDFPKPIVRSIT 301
RESULT 11
US-10-144-188-14
; Sequence 14, Application US/10144188
; Publication No. US20030170212A1
; GENERAL INFORMATION:
; APPLICANT: Cai, Zeling
; APPLICANT: Jackson, Michael R.
; APPLICANT: Peterson, Per A.
; APPLICANT: Shi, Weixing
; APPLICANT: Kong, Yan
; APPLICANT: Degraw, Juli
; TITLE OF INVENTION: Ex-Vivo Priming For Generating Cytotoxic T Lymphocytes Specific
; FILE OF INVENTION: For No. US20030170212A1-Tumor Antigens To Treat Autoimmune And
; FILE REFERENCE: PRI0010 ORT-1627
; CURRENT APPLICATION NUMBER: US/10/144,188
; CURRENT FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: 60/291,300
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide antigen
US-10-144-188-14
Query Match 77.1%; Score 108; DB 14; Length 421;
Best Local Similarity 90.9%; Pred. No. 3.9e-08;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 GYGQSIIVDRDPFKPIVRSIT 23
DB 280 GYGQCIIVDHPDFPKPIVRSIT 301
RESULT 10
US-10-144-188-14
; Sequence 14, Application US/10144188
; Publication No. US20030170212A1
; GENERAL INFORMATION:
; APPLICANT: Cai, Zeling
; APPLICANT: Jackson, Michael R.
; APPLICANT: Peterson, Per A.
; APPLICANT: Shi, Weixing
; APPLICANT: Kong, Yan
; APPLICANT: Degraw, Juli
; TITLE OF INVENTION: Ex-Vivo Priming For Generating Cytotoxic T Lymphocytes Specific
; FILE OF INVENTION: For No. US20030170212A1-Tumor Antigens To Treat Autoimmune And
; FILE REFERENCE: PRI0010 ORT-1627
; CURRENT APPLICATION NUMBER: US/10/144,188
; CURRENT FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: 60/291,300
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
; LENGTH: 421
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide antigen
US-10-144-188-14
Query Match 77.1%; Score 108; DB 14; Length 421;
Best Local Similarity 90.9%; Pred. No. 3.9e-08;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 GYGQSIIVDRDPFKPIVRSIT 23
DB 280 GYGQCIIVDHPDFPKPIVRSIT 301
RESULT 11
US-09-949-375A-19
; Sequence 19, Application US/09949375A
; Patent No. US20020172673A1
; GENERAL INFORMATION:
; APPLICANT: KLYSNER, Steen et al.
; TITLE OF INVENTION: METHOD FOR DOWN-REGULATING IGE
; FILE REFERENCE: 3631-0111P
; CURRENT APPLICATION NUMBER: US/09/949,375A
; CURRENT FILING DATE: 2002-01-18
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 432
; TYPE: PRT
; ORGANISM: mus musculus
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(432)
```

```
; FILE REFERENCE: 3631-0111P
; CURRENT APPLICATION NUMBER: US/09/949,375A
; CURRENT FILING DATE: 2002-01-18
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 25
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial amino acid sequence of SEQ ID NO: 24.
US-09-949-375A-25

Query Match
Best Local Similarity 77.1%; Score 108; DB 9; Length 332;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GYGQSIIVDRPDPFKPIVRSIT 23
Db 191 GYGQCIIVDHPDPFKPIVRSIT 212

RESULT 3
US-09-949-375A-27
; Sequence 27, Application US/09949375A
; Patent No. US20020172673A1
; GENERAL INFORMATION:
; APPLICANT: KLYSNER, Steen et al.
; TITLE OF INVENTION: METHOD FOR DOWN-REGULATING IGE
; FILE REFERENCE: 3631-0111P
; CURRENT APPLICATION NUMBER: US/09/949,375A
; CURRENT FILING DATE: 2002-01-18
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 27
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial amino acid sequence of SEQ ID NO: 26.
US-09-949-375A-27

Query Match
Best Local Similarity 77.1%; Score 108; DB 9; Length 332;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GYGQSIIVDRPDPFKPIVRSIT 23
Db 191 GYGQCIIVDHPDPFKPIVRSIT 212

RESULT 4
US-09-401-636-6
; Sequence 6, Application US/09401636
; Patent No. US20010038843A1
; GENERAL INFORMATION:
; APPLICANT: Hellman, Lars T.
; TITLE OF INVENTION: ENHANCED VACCINES
; FILE REFERENCE: 10223/006001
; CURRENT APPLICATION NUMBER: US/09/401,636
; CURRENT FILING DATE: 1999-09-22
; PRIOR APPLICATION NUMBER: US 60/106,652
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated proteins
US-09-401-636-6
```

```
Query Match
Best Local Similarity 77.1%; Score 108; DB 9; Length 341;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GYGQSIIVDRPDPFKPIVRSIT 23
Db 206 GYGQCIIVDHPDPFKPIVRSIT 227

RESULT 5
US-10-176-664-6
; Sequence 6, Application US/10176664
; Publication No. US20030031663A1
; GENERAL INFORMATION:
; APPLICANT: Hellman, Lars T.
; TITLE OF INVENTION: ENHANCED VACCINES
; FILE REFERENCE: 10223/006001
; CURRENT APPLICATION NUMBER: US/10/176,664
; CURRENT FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US/09/401,636
; PRIOR FILING DATE: 1999-09-22
; PRIOR APPLICATION NUMBER: US 60/106,652
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated proteins
US-10-176-664-6

Query Match
Best Local Similarity 77.1%; Score 108; DB 14; Length 341;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GYGQSIIVDRPDPFKPIVRSIT 23
Db 206 GYGQCIIVDHPDPFKPIVRSIT 227

RESULT 6
US-09-949-375A-20
; Sequence 20, Application US/09949375A
; Patent No. US20020172673A1
; GENERAL INFORMATION:
; APPLICANT: KLYSNER, Steen et al.
; TITLE OF INVENTION: METHOD FOR DOWN-REGULATING IGE
; FILE REFERENCE: 3631-0111P
; CURRENT APPLICATION NUMBER: US/09/949,375A
; CURRENT FILING DATE: 2002-01-18
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 20
; LENGTH: 343
; TYPE: PRT
; ORGANISM: mus musculus
; FEATURE:
; NAME/KEY: MISC FEATURE
; OTHER INFORMATION: Murine IGE heavy chain, domains C2, C3, C4, and MIGIS fragment
US-09-949-375A-20

Query Match
Best Local Similarity 77.1%; Score 108; DB 9; Length 343;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GYGQSIIVDRPDPFKPIVRSIT 23
Db 191 GYGQCIIVDHPDPFKPIVRSIT 212

RESULT 7
```



```

AAE35016
ID  AAE35016 standard; protein; 421 AA.
XX
AC  AAE35016;
XX
DT  28-MAY-2003 (first entry)
DE  Mouse immunoglobulin E (IgE).
XX
KW  Cytotoxic T lymphocyte; CTL; tumour; antigen presenting cell; allergy;
KW  lupus; autoimmune disease; rheumatoid arthritis; autoimmune nephritis;
KW  psoriasis; APC; multiple sclerosis; autoimmune thyroiditis; hay fever;
KW  inflammatory bowel disease; insulin dependent diabetes; cell therapy;
KW  Crohn's disease; allergic rhinitis; graft versus host disease; asthma;
KW  transplant rejection; mouse; immunoglobulin E; IgE.
XX
OS  Mus sp.
XX
PN  WO200292773-A2.
XX
PD  21-NOV-2002.
XX
PF  13-MAY-2002; 2002WO-US015341.
XX
PR  15-MAY-2001; 2001US-0291300P.
XX
PA  (ORTH ) ORTHO-MCNEIL PHARM INC.
XX
PI  Cai Z, Jackson MR, Peterson PA, Shi W, Kong Y, Degraw J;
XX
DR  WPI; 2003-120673/11.
XX
PT  Producing cytotoxic T lymphocytes for treating e.g. autoimmune diseases,
PT  comprises culturing CD8+ T cells with antigen presenting cells to
PT  activate precursor CD8+ T cells specific for T the cell epitopes.
XX
PS  Example 3; Col 45; 49pp; English.
XX
CC  The invention relates to a method of producing cytotoxic T lymphocytes
CC  (CTLs) specific for one or more non-tumour self antigen T cell epitopes.
CC  The method involves loading antigen presenting cells (APCs) having class
CC  I major histocompatibility complex molecules with the T cell epitopes,
CC  and culturing the CD8+ T cells with the APCs to activate precursor CD8+ T
CC  cells specific for the T cell epitopes. The invention is useful for
CC  treating autoimmune disease including rheumatoid arthritis, psoriasis,
CC  lupus, autoimmune nephritis, multiple sclerosis, autoimmune thyroiditis,
CC  insulin dependent diabetes, Crohn's disease, inflammatory bowel disease,
CC  graft versus host disease and transplant rejection and/or allergic
CC  disease such as food allergy, hay fever, allergic rhinitis, allergic
CC  asthma and venom allergy. The invention is also useful in cell therapy.
CC  The present sequence is mouse immunoglobulin E (IgE). This protein is
CC  used in the exemplification of the invention
XX
SQ  Sequence 421 AA;

Query Match      77.1%; Score 108; DB 6; Length 421;
Best Local Similarity 90.9%; Fred. No. 3.6e-08;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 GYGQSIIVDRDPFPRPIVRSIT 23
      |||||
DB      280 GYGQCIIVDHPDFPRPIVRSIT 301

Search completed: February 26, 2004, 08:23:35
Job time : 47.6 secs

```



the invention may be used as a vaccine against allergic diseases. The method of the invention is useful for inducing an immune response against autologous IGE in an animal, which is useful for downregulating autologous IGE in the animal. This method is useful in the prevention and treatment of allergic diseases such as anaphylaxis, allergic rhinitis, asthma and atopic dermatitis. The present sequence represents the DNA encoding the mouse IGE heavy chain C2-C3-C4 domain with MIGIS virus used to create the epitopes used in the method of the invention

XX Sequence 343 AA;

Query Match 77.1%; Score 108; DB 5; Length 343;  
Best Local Similarity 90.9%; Pred. No. 2.8e-08;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GYGQSIIVDRDPFKPIVRSIT 23  
DB 191 GYGQCIIVDHPDFKPIVRSIT 212

RESULT 11  
AAU80295  
ID AAU80295 standard; protein; 343 AA.

XX AC AAU80295;

XX DT 30-JUL-2002 (first entry)

XX DE Murine IGE heavy chain C2-C3-C4 with MIGIS fragment.

XX KW IGE; allergy; mouse; antiallergic; immunosuppressive; antianaphylactic;  
XX OS antihistaminic; dermatological; antiinflammatory; immunoglobulin E; IGE;  
XX KW vaccine; anaphylaxis; allergic rhinitis; asthma; atopic dermatitis;  
XX KW heavy chain C domain.

XX OS Mus sp.

XX OS Synthetic.

XX PN WO200220038-A2.

XX PD 14-MAR-2002.

XX PF 06-SEP-2001; 2001WO-DK000579.

XX PR 06-SEP-2000; 2000DK-00001326.

XX PR 15-SEP-2000; 2000US-0232831P.

XX PA (PHAR-) PHARMEXA AS.

XX PI Klysner S, Von Hoegen P, Voldborg B, Gautam A;

XX DR WPI; 2002-383033/41.

XX PT Inducing immune response against autologous immunoglobulin E in an animal, by effecting simultaneous presentation of cytotoxic T lymphocyte epitope an/or B-cell epitope derived from the immunoglobulin.

XX PS Example; Page 131-133; 15pp; English.

XX This invention relates to a novel method for inducing an immune response against autologous immunoglobulin E (IGE) in an animal. The method comprises effecting simultaneous presentation of cytotoxic T lymphocyte (CTL) epitope and/or B-cell epitope derived from IGE, and a helper cell epitope (TH epitope) which is foreign to the animal, by antigen presenting cells (APCs) of the animal's immune system. The epitopes of the invention may be used as a vaccine against allergic diseases. The method of the invention is useful for inducing an immune response against autologous IGE in an animal, which is useful for downregulating autologous IGE in the animal. This method is useful in the prevention and treatment of allergic diseases such as anaphylaxis, allergic rhinitis, asthma and atopic dermatitis. The present sequence represents the mouse IGE heavy chain C2-C3-C4 domain used to create the epitopes used in the method of the invention

XX SQ Sequence 343 AA;

Query Match 77.1%; Score 108; DB 5; Length 343;  
Best Local Similarity 90.9%; Pred. No. 2.8e-08;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GYGQSIIVDRDPFKPIVRSIT 23  
DB 191 GYGQCIIVDHPDFKPIVRSIT 212

RESULT 12  
AAE35111

ID AAE35111 standard; protein; 388 AA.

XX AC AAE35111;

XX DT 28-MAY-2003 (first entry)

XX DE Mouse immunoglobulin E (IGE) a allelic protein.

XX KW Cytotoxic T lymphocyte; CTL; tumour; antigen presenting cell; allergy;  
XX OS lupus; autoimmune disease; rheumatoid arthritis; autoimmune nephritis;  
XX KW psoriasis; APC; multiple sclerosis; autoimmune thyroiditis; hay fever;  
XX KW inflammatory bowel disease; insulin dependent diabetes; cell therapy;  
XX KW Crohn's disease; allergic rhinitis; graft versus host disease; asthma;  
XX KW transplant rejection; mouse; immunoglobulin E; IGE.

XX OS Mus sp.

XX PN WO200292773-A2.

XX PD 21-NOV-2002.

XX PF 13-MAY-2002; 2002WO-US015341.

XX PR 15-MAY-2001; 2001US-0291300P.

XX PA (ORTH ) ORTHO-MCNEIL PHARM INC.

XX PI Cai Z, Jackson WR, Peterson PA, Shi W, Kong Y, Degraw J;

XX DR WPI; 2003-120673/11.

XX PT Producing cytotoxic T lymphocytes for treating e.g. autoimmune diseases, comprising culturing CD8+ T cells with antigen presenting cells to activate precursor CD8+ T cells specific for T the cell epitopes.

XX PS Disclosure; Col 48; 49pp; English.

XX The invention relates to a method of producing cytotoxic T lymphocytes (CTLs) specific for one or more non-tumour self antigen T cell epitopes. The method involves loading antigen presenting cells (APCs) having class I major histocompatibility complex molecules with the T cell epitopes, and culturing the CD8+ T cells with the APCs to activate precursor CD8+ T cells specific for the T cell epitopes. The invention is useful for treating autoimmune disease including rheumatoid arthritis, psoriasis, lupus, autoimmune nephritis, multiple sclerosis, autoimmune thyroiditis, insulin dependent diabetes, Crohn's disease, inflammatory bowel disease, graft versus host disease and transplant rejection and/or allergic disease such as food allergy, hay fever, allergic rhinitis, allergic asthma and venom allergy. The invention is also useful in cell therapy. The present sequence is mouse immunoglobulin E (IGE) a allelic protein. This protein is used in the exemplification of the invention

XX SQ Sequence 388 AA;

Query Match 77.1%; Score 108; DB 6; Length 388;  
Best Local Similarity 90.9%; Pred. No. 3.3e-08;  
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GYGQSIIVDRDPFKPIVRSIT 23



DR WPI; 2002-383033/41.  
 DR N-PSDB; ABX51142.  
 XX Inducing immune response against autologous immunoglobulin E in an  
 PT animal, by effecting simultaneous presentation of cytotoxic T lymphocyte  
 PT epitope an/or B-cell epitope derived from the immunoglobulin.  
 XX  
 XX Disclosure; Page 144-145; 151pp; English.  
 XX  
 CC This invention relates to a novel method for inducing an immune response  
 CC against autologous immunoglobulin E (IgE) in an animal. The method  
 CC comprises effecting simultaneous presentation of cytotoxic T lymphocyte  
 CC (CTL) epitope and/or B-cell epitope derived from IgE, and T helper cell  
 CC epitope (TH epitope) which is foreign to the animal, by antigen  
 CC presenting cells (APCs) of the animal's immune system. The epitopes of  
 CC the invention may be used as a vaccine against allergic diseases. The  
 CC method of the invention is useful for inducing an immune response against  
 CC autologous IgE in an animal, which is useful for downregulating  
 CC autologous IgE in the animal. This method is useful in the prevention and  
 CC treatment of allergic diseases such as anaphylaxis, allergic rhinitis,  
 CC asthma and atopic dermatitis. The present sequence represents the mouse  
 CC IgE heavy chain C2-C3-C4 domain optimised for an E.Coli expression system  
 CC used to create the epitopes used in the method of the invention  
 XX  
 XX Sequence 332 AA;  
 SQ

Query Match 77.1%; Score 108; DB 5; Length 332;  
 Best Local Similarity 90.9%; Pred. No. 2.7e-08;  
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 GYGQSIIVDRDPDPKPIVRSIT 23  
 DB 191 GYGQCIIVDHPDPKPIVRSIT 212

RESULT 9  
 AAU80297  
 ID AAU80297 standard; protein; 332 AA.  
 AC AAU80297;  
 XX 30-JUL-2002 (first entry)  
 DE Mouse IgE heavy chain C2-C3-C4 domain with MIGIS fragment.  
 XX  
 KW IgE; allergy; mouse; antiallergic; immunosuppressive; antianaphylactic;  
 KW antiasthmatic; dermatological; antiinflammatory; immunoglobulin E; IgE;  
 KW vaccine; anaphylaxis; allergic rhinitis; asthma; atopic dermatitis;  
 KW heavy chain C domain; MIGIS.  
 XX  
 OS Mus sp.  
 OS Synthetic.  
 XX WO200220038-A2.  
 XX 14-MAR-2002.  
 XX 06-SEP-2001; 2001WO-DK000579.  
 XX 06-SEP-2000; 2000DK-00001326.  
 XX 15-SEP-2000; 2000US-0232831P.  
 XX (PHAR-) PHARMEXA AS.  
 XX Klysner S, Von Hoegen P, Voldborg B, Gautam A;  
 XX WPI; 2002-383033/41.  
 XX 06-SEP-2001; 2001WO-DK000579.  
 XX 06-SEP-2000; 2000DK-00001326.  
 XX 15-SEP-2000; 2000US-0232831P.  
 XX (PHAR-) PHARMEXA AS.  
 XX Klysner S, Von Hoegen P, Voldborg B, Gautam A;  
 XX WPI; 2002-383033/41.  
 XX Inducing immune response against autologous immunoglobulin E in an  
 PT animal, by effecting simultaneous presentation of cytotoxic T lymphocyte  
 PT epitope an/or B-cell epitope derived from the immunoglobulin.  
 XX  
 XX Example; Page 137-138; 151pp; English.

XX This invention relates to a novel method for inducing an immune response  
 CC against autologous immunoglobulin E (IgE) in an animal. The method  
 CC comprises effecting simultaneous presentation of cytotoxic T lymphocyte  
 CC (CTL) epitope and/or B-cell epitope derived from IgE, and T helper cell  
 CC epitope (TH epitope) which is foreign to the animal, by antigen  
 CC presenting cells (APCs) of the animal's immune system. The epitopes of  
 CC the invention may be used as a vaccine against allergic diseases. The  
 CC method of the invention is useful for inducing an immune response against  
 CC autologous IgE in an animal, which is useful for downregulating  
 CC autologous IgE in the animal. This method is useful in the prevention and  
 CC treatment of allergic diseases such as anaphylaxis, allergic rhinitis,  
 CC asthma and atopic dermatitis. The present sequence represents the mouse  
 CC IgE heavy chain C2-C3-C4 domain with the MIGIS fragment used to create  
 CC the epitopes used in the method of the invention  
 XX  
 XX Sequence 332 AA;  
 SQ

Query Match 77.1%; Score 108; DB 5; Length 332;  
 Best Local Similarity 90.9%; Pred. No. 2.7e-08;  
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 GYGQSIIVDRDPDPKPIVRSIT 23  
 DB 191 GYGQCIIVDHPDPKPIVRSIT 212

RESULT 10  
 AAU80296  
 ID AAU80296 standard; protein; 343 AA.  
 AC AAU80296;  
 XX 30-JUL-2002 (first entry)  
 DE Murine IgE heavy chain C2-C3-C4 with MIGIS fragment.  
 XX  
 KW IgE; allergy; mouse; antiallergic; immunosuppressive; antianaphylactic;  
 KW antiasthmatic; dermatological; antiinflammatory; immunoglobulin E; IgE;  
 KW vaccine; anaphylaxis; allergic rhinitis; asthma; atopic dermatitis;  
 KW heavy chain C domain.  
 XX  
 OS Mus sp.  
 OS Synthetic.  
 XX WO200220038-A2.  
 XX 14-MAR-2002.  
 XX 06-SEP-2001; 2001WO-DK000579.  
 XX 06-SEP-2000; 2000DK-00001326.  
 XX 15-SEP-2000; 2000US-0232831P.  
 XX (PHAR-) PHARMEXA AS.  
 XX Klysner S, Von Hoegen P, Voldborg B, Gautam A;  
 XX WPI; 2002-383033/41.  
 XX N-PSDB; ABX51140.  
 XX Inducing immune response against autologous immunoglobulin E in an  
 PT animal, by effecting simultaneous presentation of cytotoxic T lymphocyte  
 PT epitope an/or B-cell epitope derived from the immunoglobulin.  
 XX  
 XX Example; Page 135-136; 151pp; English.  
 XX  
 CC This invention relates to a novel method for inducing an immune response  
 CC against autologous immunoglobulin E (IgE) in an animal. The method  
 CC comprises effecting simultaneous presentation of cytotoxic T lymphocyte  
 CC (CTL) epitope and/or B-cell epitope derived from IgE, and T helper cell  
 CC epitope (TH epitope) which is foreign to the animal, by antigen  
 CC presenting cells (APCs) of the animal's immune system. The epitopes of



RESULT 4  
AAE35112  
ID AAE35112 standard; protein; 423 AA.  
XX  
AC AAE35112;  
XX  
DT 28-MAY-2003 (first entry)  
XX  
DE Mouse immunoglobulin E (IgE) b allelic protein.  
XX  
KW Cytotoxic T lymphocyte; CTL; tumour; antigen presenting cell; allergy;  
KW lupus; autoimmune disease; rheumatoid arthritis; autoimmune nephritis;  
KW psoriasis; APC; multiple sclerosis; autoimmune thyroiditis; hay fever;  
KW inflammatory bowel disease; insulin dependent diabetes; cell therapy;  
KW Crohn's disease; allergic rhinitis; graft versus host disease; asthma;  
KW transplant rejection; mouse; immunoglobulin E; IgE.  
XX  
OS Mus sp.  
XX  
FN WO200292773-A2.  
XX  
XX  
PD 21-NOV-2002.  
XX  
PF 13-MAY-2002; 2002WO-US015341.  
XX  
PR 15-MAY-2001; 2001US-0291300P.  
XX  
XX (ORTH ) ORTHO-MCNEIL PHARM INC.  
XX  
XX Cai Z, Jackson MR, Peterson PA, Shi W, Kong Y, Degraw J;  
XX WPI; 2003-120673/11.  
XX  
XX Producing cytotoxic T lymphocytes for treating e.g. autoimmune diseases.  
XX comprises culturing CD8+ T cells with antigen presenting cells to  
XX activate precursor CD8+ T cells specific for T the cell epitopes.  
XX  
XX Disclosure; Fig 1; 49pp; English.  
XX  
XX The invention relates to a method of producing cytotoxic T lymphocytes  
XX (CTLs) specific for one or more non-tumour self antigen T cell epitopes.  
XX The method involves loading antigen presenting cells (APCs) having class  
XX I major histocompatibility complex molecules with the T cell epitopes,  
XX and culturing the CD8+ T cells with the APCs to activate precursor CD8+ T  
XX cells specific for the T cell epitopes. The invention is useful for  
XX treating autoimmune disease including rheumatoid arthritis, psoriasis,  
XX lupus, autoimmune nephritis, multiple sclerosis, autoimmune thyroiditis,  
XX insulin dependent diabetes, Crohn's disease, inflammatory bowel disease,  
XX graft versus host disease and transplant rejection and/or allergic  
XX disease such as food allergy, hay fever, allergic rhinitis, allergic  
XX asthma and venom allergy. The invention is also useful in cell therapy.  
XX The present sequence is mouse immunoglobulin E (IgE) b allelic protein.  
XX This protein is used in the exemplification of the invention  
XX  
SQ Sequence 423 AA;  
Query Match 82.9%; Score 116; DB 6; Length 423;  
Best Local Similarity 91.3%; Pred. No. 2e-09;  
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 2 GYGQSVDRDPDFPKPIVRSITL 24  
Db 281 GYGQCVVDRDPDFPKPIVRSITL 303  
RESULT 5  
AAV17415  
ID AAV17415 standard; protein; 561 AA.  
XX  
AC AAV17415;  
XX

DT 26-JUL-1999 (first entry)  
XX  
DE Mouse immunoglobulin E heavy chain.  
XX  
KW Mouse; immunoglobulin E; IgE; allergy; transgenic animal.  
XX  
OS Mus sp.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..19  
FT Protein 20..561  
FT /label= signal  
FT /label= immunoglobulin\_E\_heavy\_chain  
XX  
PN EP921189-A1.  
XX  
PD 09-JUN-1999.  
XX  
PF 13-NOV-1998; 98EP-00309340.  
XX  
PR 14-NOV-1997; 97JP-00313989.  
XX  
PA (SANY ) SANKYO CO LTD.  
PA (TOKM-) TOKYO METROPOLITAN INST MEDICAL SCI.  
XX  
PI Karasuyama H, Yonemawa H, Taya C, Matsuoka K;  
XX WPI; 1999-315404/27.  
DR N-PSDB; AAX61084.  
XX  
PT Transgenic non-human animal allergy models.  
XX  
PS Claim 27; Page 24-27; 42pp; English.  
XX  
CC The present invention describes a transgenic, non-human animal with its  
CC genome altered to constitutively express a molecule having a constant  
CC region which can bind an IGE receptor on mast cells in the animals, the  
CC molecule having an immunoglobulin structure and being further capable of  
CC specifically binding a predetermined antigen. The transgenic animal is  
CC useful as a model for evaluating the activity and the ability of  
CC substances i.e. with antiallergic activity, to affect any allergic  
CC reaction caused in the animal by the administration of the antigen to the  
CC animal, and applying the substance to be evaluated. The present sequence  
CC is the mouse immunoglobulin E (IGE) heavy chain, given in the present  
CC invention  
XX  
SQ Sequence 561 AA;  
Query Match 82.9%; Score 116; DB 2; Length 561;  
Best Local Similarity 91.3%; Pred. No. 2.8e-09;  
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 2 GYGQSVDRDPDFPKPIVRSITL 24  
Db 419 GYGQCVVDRDPDFPKPIVRSITL 441  
RESULT 6  
AAV79997  
ID AAV79997 standard; protein; 313 AA.  
XX  
AC AAV79997;  
XX  
DT 15-MAY-2000 (first entry)  
XX  
DE Mouse immunoglobulin E epsilon heavy chain SEQ ID NO:4.  
XX  
KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;  
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;  
KW antibody; allergy; allergic disease; immunisation; anti-allergic;  
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.  
XX  
OS Mus sp.

CC (functional in genetically diverse subjects), in addition to a B cell  
 CC target epitope, have increased immunogenicity and may include cyclic  
 CC constraints (disulfide bridge) to stabilise conformational features and  
 CC maximize cross-reactivity to the natural target. They induce safe (non-  
 CC anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino acid  
 CC sequences used in the exemplification of the present invention  
 XX  
 SQ Sequence 25 AA;

Query Match 100.0%; Score 140; DB 3; Length 25;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-14; Indels 0; Gaps 0;  
 Matches 25; Conservative 0; Mismatches 0;

QY 1 CGYGYQSIVDRPDPFKPIVRSITLC 25  
 |||||  
 DB 1 CGYGYQSIVDRPDPFKPIVRSITLC 25

## RESULT 2

AAAY80018

ID AAY80018 standard; peptide; 45 AA.  
 XX  
 AC AAY80018;  
 XX  
 DT 15-MAY-2000 (first entry)  
 XX  
 DE IGE immunogenic peptide conjugate SEQ ID NO:25.

XX Immunoglobulin E; IGE; epsilon heavy chain; antigenic; antigen;  
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;  
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;  
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.  
 XX  
 OS Unidentified.  
 XX  
 PN WO9967293-A1.  
 XX  
 PD 29-DEC-1999.  
 XX  
 PF 21-JUN-1999; 99WO-US013959.  
 XX  
 PR 20-JUN-1998; 98US-00100287.  
 XX  
 PA (UNBI-) UNITED BIOMEDICAL INC.

PI Wang CY, Walfield AM;  
 XX  
 DR WPI; 2000-160578/14.  
 XX  
 PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions  
 XX for immunization against allergy.

PS Claim 14; Page 76; 155pp; English.  
 XX  
 CC The present invention describes immunoglobulin E (IGE)-CH3 domain  
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic and  
 CC anti-asthmatic properties. (I) induces polyclonal antibodies specific for  
 CC a target effector site on the epsilon-heavy chain of IGE, and so  
 CC preventing triggering and activation of mast cells and basophils and  
 CC downregulation of IGE synthesis. Conjugates, or fusion peptides,  
 CC containing (I) are used for active immunisation against IGE-mediated  
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy  
 CC dermatitis. Nucleic acids that encode these compounds are useful for  
 CC recombinant production of corresponding peptides or in DNA vaccines.  
 CC Conjugates of (I) that include a promiscuous T helper cell epitope  
 CC (functional in genetically diverse subjects), in addition to a B cell  
 CC target epitope, have increased immunogenicity and may include cyclic  
 CC constraints (disulfide bridge) to stabilise conformational features and  
 CC maximize cross-reactivity to the natural target. They induce safe (non-  
 CC anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino acid  
 XX sequences used in the exemplification of the present invention  
 SQ Sequence 45 AA;

Query Match 92.1%; Score 129; DB 3; Length 45;  
 Best Local Similarity 92.0%; Pred. No. 1.6e-12;  
 Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGYGYQSIVDRPDPFKPIVRSITLC 25  
 |||||  
 DB 21 CGYGYQSIVDRPDPFKPIVRSITLC 45

## RESULT 3

AAAY80017

ID AAY80017 standard; peptide; 46 AA.  
 XX  
 AC AAY80017;  
 XX  
 DT 15-MAY-2000 (first entry)  
 XX  
 DE IGE immunogenic peptide conjugate SEQ ID NO:24.

XX Immunoglobulin E; IGE; epsilon heavy chain; antigenic; antigen;  
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;  
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;  
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX OS Unidentified.  
 XX  
 PN WO9967293-A1.  
 XX  
 PD 29-DEC-1999.  
 XX  
 PF 21-JUN-1999; 99WO-US013959.  
 XX  
 PR 20-JUN-1998; 98US-00100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

PI Wang CY, Walfield AM;  
 XX  
 DR WPI; 2000-160578/14.  
 XX  
 PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions  
 XX for immunization against allergy.

PS Claim 14; Page 76; 155pp; English.

XX The present invention describes immunoglobulin E (IGE)-CH3 domain  
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic and  
 CC anti-asthmatic properties. (I) induces polyclonal antibodies specific for  
 CC a target effector site on the epsilon-heavy chain of IGE, and so  
 CC preventing triggering and activation of mast cells and basophils and  
 CC downregulation of IGE synthesis. Conjugates, or fusion peptides,  
 CC containing (I) are used for active immunisation against IGE-mediated  
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy  
 CC dermatitis. Nucleic acids that encode these compounds are useful for  
 CC recombinant production of corresponding peptides or in DNA vaccines.  
 CC Conjugates of (I) that include a promiscuous T helper cell epitope  
 CC (functional in genetically diverse subjects), in addition to a B cell  
 CC target epitope, have increased immunogenicity and may include cyclic  
 CC constraints (disulfide bridge) to stabilise conformational features and  
 CC maximize cross-reactivity to the natural target. They induce safe (non-  
 CC anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino acid  
 XX sequences used in the exemplification of the present invention  
 SQ Sequence 46 AA;

Query Match 92.1%; Score 129; DB 3; Length 46;  
 Best Local Similarity 92.0%; Pred. No. 1.6e-12;  
 Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGYGYQSIVDRPDPFKPIVRSITLC 25  
 |||||  
 DB 22 CGYGYQSIVDRPDPFKPIVRSITLC 46

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 08:16:50 / Search time 46.6 Seconds  
(without alignments)  
151.581 Million cell updates/sec

Title: US-09-701-623c-8

Perfect score: 140

Sequence: 1 CGYQSYIVDRDPFKPIVRSITLC 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*

1: Geneseqp1980s.\*

2: Geneseqp1990s.\*

3: Geneseqp2000s.\*

4: Geneseqp2001s.\*

5: Geneseqp2002s.\*

6: Geneseqp2003as.\*

7: Geneseqp2003bs.\*

8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	140	100.0	25	3 AAY80001	Aay80001 Optimised
2	129	92.1	45	3 AAY80018	Aay80018 IGE immu
3	129	92.1	45	3 AAY80017	Aay80017 IGE immu
4	116	82.9	423	6 AAE35112	Rae35112 Mouse imm
5	116	82.9	561	2 AAY17415	Aay17415 Mouse imm
6	113	80.7	313	3 AAY79997	Aay79997 Mouse imm
7	108	77.1	332	5 AAU80298	Aau80298 Murine Ig
8	108	77.1	332	5 AAU80299	Aau80299 Murine Ig
9	108	77.1	332	5 AAU80297	Aau80297 Murine Ig
10	108	77.1	343	5 AAU80296	Aau80296 Murine Ig
11	108	77.1	343	5 AAU80295	Aau80295 Murine Ig
12	108	77.1	388	6 AAE35111	Rae35111 Mouse imm
13	108	77.1	421	5 AAU80300	Aau80300 Mouse IGE
14	108	77.1	421	6 ABP96586	Abp96586 Mouse IGE
15	108	77.1	421	6 AAE35016	Rae35016 Mouse imm
16	108	77.1	432	5 AAU80294	Aau80294 Murine Ig
17	106	75.7	25	3 AAY80000	Aay80000 Optimised
18	85	60.7	114	6 ABG74776	Abg74776 Rat Ige C
19	85	60.7	313	3 AAY79996	Aay79996 Rat immu
20	85	60.7	340	3 AAB03643	Aab03643 Rat Ige h
21	85	60.7	341	3 AAB06206	Rab06206 Immunogen
22	85	60.7	346	6 ABG74785	Abg74785 Human CH2
23	85	60.7	428	6 ABP96589	Abp96589 Rat Ige h
24	74	52.9	25	3 AAY80077	Aay80077 Optimised
25	66	47.1	25	3 AAY79999	Aay79999 Optimised

ALIGNMENTS				
RESULT 1				
AAY80001				
ID	AAY80001	standard; peptide; 25 AA.		
XX	XX			
AC	AAY80001;			
XX	XX			
DT	15-MAY-2000	(first entry)		
XX	XX			
DE	Optimised IGE-CH3 domain antigen peptide for mouse IGE.			
XX	XX			
KW	Immunoglobulin E; IGE; epsilon heavy chain; antigenic; antigen;			
KW	immunogenic; immunostimulatory; carrier protein; helper T cell epitope;			
KW	antibody; allergy; allergic disease; immunisation; anti-allergic;			
KW	anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.			
XX	XX			
OS	Mus sp.			
OS	Synthetic.			
XX	XX			
PN	WO9967293-A1.			
XX	XX			
PD	29-DEC-1999.			
XX	XX			
PF	21-JUN-1999;	99WO-US013959.		
XX	XX			
PR	20-JUN-1998;	98US-00100287.		
XX	XX			
PA	(UNBI-) UNITED BIOMEDICAL INC.			
XX	XX			
PI	Wang CY, Walfield AM;			
XX	XX			
DR	WPI; 2000-160578/14.			
XX	XX			
PT	New antigenic peptide from the CH3 domain of immunoglobulin E, fusions			
PT	for immunization against allergy.			
XX	XX			
PS	Claim 1; Page 100; 155pp; English.			
XX	XX			
CC	The present invention describes immunoglobulin E (IGE)-CH3 domain antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic and anti-asthmatic properties. (I) induces polyclonal antibodies specific for a target effector site on the epsilon-heavy chain of IGE, and so preventing triggering and activation of mast cells and basophils and downregulation of IGE synthesis. Conjugates, or fusion peptides, containing (I) are used for active immunisation against IGE-mediated allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy dermatitis. Nucleic acids that encode these compounds are useful for recombinant production of corresponding peptides or in DNA vaccines. Conjugates of (I) that include a promiscuous T helper cell epitope			

```

QY 1 CGYGYQSIVDRPDPFK 16
Db 47 CGYGYLGILMLPLPK 62

RESULT 14
Q88MD1
ID Q98MD1 PRELIMINARY; PRT; 753 AA.
AC Q98MD1;
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE DNA uptake protein.
GN MLO627.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asanizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP002995; BAB48182.1;
DR InterPro; IPR004477; COMEC_N-term.
DR Pfam; PF03772; Competence; 1.
DR TIGRFAMs; TIGR00360; COMEC_N-term; 1.
KW Complete proteome.
SQ SEQUENCE 753 AA; 80233 MW; C416697517B91EA9 CRC64;

Query Match 34.3%; Score 48; DB 16; Length 753;
Best Local Similarity 53.3%; Pred. No. 1e+02;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 GYQSIVDRPDPFKPI 18
Db 44 GYSLAVEPDPFAKPV 58

RESULT 15
Q81283
ID Q81283 PRELIMINARY; PRT; 832 AA.
AC Q81283;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein.
GN PFA0285C.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22255708; PubMed=12368867;
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
RA Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
RA Humphray S., Jagels K., James K.D., Johnson D., Kernhorn A.,
RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
RA Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,
RA Oliver K., Ormond D., Price C., Quail M.A., Rabinowitsch E.,
RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,

```

```

RA Seeger K., Sharp S., Smith P., Squares R., Squares S., Stevens K.,
RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
RA Sulston J.E., Craig A., Newbold C., Barrell B.G.;
RT "Sequence of plasmodium falciparum chromosomes 1, 3-9 and 13.";
RL Nature 419:527-531(2002).
DR EMBL; AL031745; CAD49017.1; -.
KW Hypothetical protein.
SQ SEQUENCE 832 AA; 99310 MW; 3E4C1249C9FD459B CRC64;

Query Match 34.3%; Score 48; DB 5; Length 832;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 YOSIVDRPDPFK 16
Db 696 YBSIIDAPDFSK 707

Search completed: February 26, 2004, 08:27:21
Job time : 35.4 secs

```

Db 177 CQGYVLADNRSCVCSNHFVTERQGLSSPDPFPYPRAL 219

## RESULT 11

```
Q8AW90 PRELIMINARY; PRT; 722 AA.
AC Q8AW90;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Mannose-binding lectin-associated serine protease.
GN MASP.
OS Lampetra japonica (Japanese lamprey) (Entosphenus japonicus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Lethenteron.
OX NCBI_TaxID=94589;
RN [1]
RP SEQUENCE FROM N.A.
RA Endo Y., Nonaka M., Saiga H., Kakinuma Y., Takahashi M.,
RA Matsumita M., Fujita T.;
RT "Ancient origin and extensive distribution of mannose-binding lectin-
RT associated serine protease-3 in vertebrate lineage.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB078894; BAC41492.1; JOINED.
DR EMBL; AB078887; BAC41492.1; JOINED.
DR EMBL; AB078888; BAC41492.1; JOINED.
DR EMBL; AB078889; BAC41492.1; JOINED.
DR EMBL; AB078890; BAC41492.1; JOINED.
DR EMBL; AB078891; BAC41492.1; JOINED.
DR EMBL; AB078892; BAC41492.1; JOINED.
DR EMBL; AB078893; BAC41492.1; JOINED.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0005523; F:sugar binding; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0007157; P:heterophilic cell adhesion; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000152; ASX_hydroxy_S.
DR InterPro; IPR000859; CUB.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR01881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR00436; Sushi_SCR_CCP.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00084; sushi; 2.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00032; CCP; 2.
DR SMART; SM00042; CUB; 2.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Lectin; Protease.
SQ SEQUENCE 722 AA; 79655 MW; BAEEBC3579F020B42 CRC64;
```

Query Match 34.6%; Score 48.5; DB 13; Length 722;  
Best Local Similarity 25.6%; Pred. No. 83;  
Matches 11; Conservative 4; Mismatches 7; Indels 21; Gaps 1;

QY 1 CQGY-----QSIVDRPDPFKPIVRSITL 22

Db 177 CQGYVLADNRSCVCSNHFVTERQGLSSPDPFPYPRAL 219

## RESULT 12

```
Q8X227 PRELIMINARY; PRT; 960 AA.
AC Q8X227;
DT 01-MAR-2002 (TREMELrel. 20, Created)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE O-acetyltransferase.
GN CAS1.
OS Cryptococcus neoformans var. neoformans.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX NCBI_TaxID=40410;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21561249; PubMed=11703667;
RA Janbon G., Himmelreich U., Moyrand F., Improvisi L., Dromer F.;
RT "Janbon is a membrane protein necessary for the O-acetylation of the
RT Cryptococcus neoformans capsular polysaccharide.";
RL Mol. Microbiol. 42:453-467(2001).
DR EMBL; AF355592; AAL35099.1; -.
DR GO; GO:0016740; F:transferase activity; IEA.
DR KW Transferase.
SQ SEQUENCE 960 AA; 107503 MW; 6AA05D5FCD2650F7 CRC64;
```

Query Match 34.6%; Score 48.5; DB 3; Length 960;  
Best Local Similarity 41.4%; Pred. No. 11e+02;  
Matches 12; Conservative 4; Mismatches 6; Indels 7; Gaps 2;

QY 1 CQGY-----QSIVDRPDPFKPIVRSITL 24

Db 668 CAYGFVKFNEHQISDRPWF--PVMRTATL 694

## RESULT 13

```
Q81R78 PRELIMINARY; PRT; 292 AA.
AC Q81R78;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Hypothetical protein.
GN BA2178.
OS Bacillus anthracis (strain Ames).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=198094;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=22608414; PubMed=12721629;
RA Read T.D., Peterson S.N., Tourasse N., Baillie L.W., Paulsen I.T.,
RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
RA Holtzapple E.K., Okstad O.A., Helgason E., Rillstone J., Wu M.,
RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.,
RA DeBoy R.T., Madupu R., Daugherty S.C., Durkin A.S., Haft D.H.,
RA Nelson W.C., Peterson J.D., Pop M., Khouli H.M., Radune D.,
RA Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,
RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,
RA Hazen A., Cline R., Redmond C., Thwaite J.E., White O., Salzberg S.L.,
RA Thomson B., Friedlander A.M., Koehler T.M., Hanna P.C., Kolsto A.-B.,
RA Fraser C.M.;
RT "The genome sequence of Bacillus anthracis Ames and comparison to
RT closely related bacteria.";
RL Nature 423:81-86(2003).
DR EMBL; AE017030; AAP26057.1; -.
DR TIGR; BA2178; -.
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. .; IEA.
DR InterPro; IPR000051; SAM_bind.
DR KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 292 AA; 33417 MW; 2594F003A726E29B CRC64;
```

Query Match 34.3%; Score 48; DB 16; Length 292;  
Best Local Similarity 50.0%; Pred. No. 38;  
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Best Local Similarity 35.3%; Pred. No. 40;  
Matches 6; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 6 QSVDRPDPFKPIVRSI 22  
DB 117 EAAQRPLPEPLRLQ 133

## RESULT 9

ID Q9V917 PRELIMINARY; PRT; 1449 AA.  
AC Q9V917;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
DE CG16868 protein.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkeley;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Anagnostou P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers J.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Efankoch C., Baldwin D.,  
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berhan B.P., Bhandari D., Bolshakov S.,  
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,  
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Hartis N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reiert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster."  
RL Science 287:2185-2195(2000).  
DR EMBL; AF003792; RAF57485.1;  
DR FlyBase; FBgn0034498; CG16868.  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR004010; Cache.  
DR InterPro; IPR020235; VWP\_A.  
DR Pfam; PF02743; VWA; 1.  
DR SMART; SM00327; VWA; 1.  
DR PROSITE; PS02334; VWFA; 1.  
SQ SEQUENCE 1449 AA; 162090 MW; 70E3A155C4185D74 CRC64;

Query Match 35.0%; Score 49; DB 5; Length 1449;  
Best Local Similarity 33.3%; Pred. No. 14e+02;  
Matches 10; Conservative 6; Mismatches 6; Indels 8; Gaps 1;

QY 2 GYGYQSVDR-----PDFPKPIVRSIT 23  
DB 609 GLGYAFLLDRSTGNTLAHPAPRPLIQRET 638

## RESULT 10

ID Q9PSZ5 PRELIMINARY; PRT; 722 AA.  
AC Q9PSZ5;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Mannose-binding lectin-associated serine protease.  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;  
OC Petromyzontiformes; Petromyzontidae; Lethenteron.  
OX NCBI\_TaxID=94989;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=99008558; PubMed=9794427;  
RA Endo Y., Takahashi M., Nakao M., Saiga H., Sekine H., Matsushita M.,  
RA Nonaka M., Fujita T.;  
RT "Two lineages of mannose-binding lectin-associated serine protease  
(MASP) in vertebrates."  
RL J. Immunol. 161:4924-4930(1998).  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.  
DR EMBL; AB009075; BAA86868.1; -;  
DR HSP; P00763; IDPO.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0005529; F:sugar binding; IEA.  
DR GO; GO:0004235; F:trypsin activity; IEA.  
DR GO; GO:0007157; P:inotrophic cell adhesion; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR00152; Asx\_hydroxyl\_S.  
DR InterPro; IPR000859; CUB.  
DR InterPro; IPR009003; Cys Ser trypsin.  
DR InterPro; IPR001881; EGF\_Ca\_-like.  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR001254; Peptidase S1.  
DR InterPro; IPR001314; Peptidase S1A.  
DR Pfam; PF00431; CUB; 2.  
DR Pfam; PF00084; sushi; 2.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00032; CCP; 2.  
DR SMART; SM00042; CUB; 2.  
DR SMART; SM00179; EGF\_CA; 1.  
DR SMART; SM00020; Trp\_SPC; 1.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 1.  
DR PROSITE; PS01180; CUB; 2.  
DR PROSITE; PS01186; EGF\_2; 1.  
DR PROSITE; PS01187; EGF\_CA; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
DR EGF-like domain; Hydrolase; Lectin; Protease; Serine protease.  
SQ SEQUENCE 722 AA; 79591 MW; 81DEB4F78F540B20 CRC64;

Query Match 34.6%; Score 48.5; DB 13; Length 722;  
Best Local Similarity 25.8%; Pred. No. 83;  
Matches 11; Conservative 4; Mismatches 7; Indels 21; Gaps 1;

QY 1 CGYGY-----QSVDRPDPFKPIVRSI 22  
| : ||| : |  
| : ||| : |





```

DR GO: 00016787; F.hydrolase activity; IEA.
DR GO: 0009231; P.vitamin B2 biosynthesis; IEA.
DR InterPro: IPR00926; GTP_cyclohydrol.
DR Pfam: PF00925; GTP_cycloHydro2; 1.
DR Hydrolase.
KW NON_TER
FT
SQ .SEQUENCE 164 AA; 18687 MW; A3BD701D82D0A700 CRC64;
      1
Query Match          35.7%; Score 50; DB 2; Length 164;
Best Local Similarity 38.1%; Pred. No. 9.9;
Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 5 YQSIVDRPDFPKPIVRSITLC 25
   ||::||:|::|:|
DB 3 YAIIEPNPYSNFIIRHSC 23

RESULT 4
QB14C0 PRELIMINARY; PRT; 338 AA.
ID QB14C0 AC QB14C0
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE DE Hypothetical protein Y69H2.14.
GN Y69H2.14.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabdittidae; Pelodierinae; Caenorhabditis.
OC NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA McMurray A.A.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RX MEDLINE=99089633; PubMed=9851916;
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99089633; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z98877; CAD56615.1; -.
DR Wormpep; Y69H2.14; CE32542.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR002486; Col_cuticle_N.
DR Pfam; PF01391; Collagen_3.
DR Pfam; PF01484; Col_cuticle_N; 1.
KW Hypothetical protein.
SQ SEQUENCE 338 AA; 33387 MW; 99205E0DEFA3CFB8 CRC64;
      1
Query Match          35.7%; Score 50; DB 5; Length 338;
Best Local Similarity 44.4%; Pred. No. 21;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 YGQSVIVDRPDFPKPIVR 20
   ||::||:|::|
DB 113 YGHQGYAQRFQPQPVYR 130

RESULT 5
QB14C0 PRELIMINARY; PRT; 443 AA.
ID QB14C0 AC QB14C0
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE B1103C09.36 protein.
GN B1103C09.36.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartioideae; Oryzeae; Oryza.
OC NCBI_TaxID=33947;

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 08:18:11 ; Search time 31.4 seconds  
(without alignments)

251.209 Million cell updates/sec

Title: US-09-701-623c-8

Perfect score: 140

Sequence: 1 CGYGYQSIWDRPDPKPIVRSITLC 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25: \*  
1: sp\_archaea: \*  
2: sp\_bacteria: \*  
3: sp\_fungi: \*  
4: sp\_human: \*  
5: sp\_invertebrate: \*  
6: sp\_mammal: \*  
7: sp\_mhc: \*  
8: sp\_organelle: \*  
9: sp\_phage: \*  
10: sp\_plant: \*  
11: sp\_rodent: \*  
12: sp\_virus: \*  
13: sp\_vertebrate: \*  
14: sp\_unclassified: \*  
15: sp\_rviro: \*  
16: sp\_bacteriap: \*  
17: sp\_archaeap: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52.5	37.5	460	10 Q94D33	Q94D33 oryza sativ
2	51.5	36.8	186	5 Q77123	Q77123 urechis cau
3	50	35.7	164	2 Q8RPM4	Q8RPM4 ehrlichia c
4	50	35.7	338	5 Q814C0	Q814C0 caenorhabd
5	49.5	35.4	443	10 Q8LQV5	Q8LQV5 oryza sativ
6	49	35.0	165	1 Q93701	Q93701 sulfolobus
7	49	35.0	232	16 Q8YKE8	Q8YKE8 anabaena sp
8	49	35.0	433	15 Q9RTJ1	Q9RTJ1 deinococcus
9	49	35.0	1449	5 Q9V917	Q9V917 drosophila
10	48.5	34.6	722	13 Q9PSZ5	Q9PSZ5 lampetra ja
11	48.5	34.6	722	13 Q8AW90	Q8AW90 lampetra ja
12	48.5	34.6	960	3 Q8X227	Q8X227 cryptococcu
13	48	34.3	292	16 Q81R78	Q81R78 bacillus an
14	48	34.3	753	16 Q98MD1	Q98MD1 rhizobium l
15	48	34.3	832	5 Q81283	Q81283 plasmodium
16	48	34.3	977	13 Q91925	Q91925 xenopus lae

17	48	34.3	1259	10 Q40983	Q40983 pisum sativ
18	48	34.3	1545	10 Q7XL81	Q7XL81 oryza sativ
19	47	33.6	219	16 Q7W378	Q7W378 bordetella
20	47	33.6	219	16 Q7W4Q2	Q7W4Q2 bordetella
21	47	33.6	219	16 Q7VU73	Q7VU73 bordetella
22	47	33.6	352	16 Q7V4D3	Q7V4D3 prochloroco
23	47	33.6	534	12 Q6S856	Q6S856 beet yellow
24	47	33.6	591	16 Q92UW6	Q92UW6 rhizobium m
25	47	33.6	914	16 Q8ZDL2	Q8ZDL2 yerseinia pe
26	46.5	33.2	181	5 Q7YZE4	Q7YZE4 drosophila
27	46.5	33.2	182	5 Q9VIO8	Q9VIO8 drosophila
28	46.5	33.2	182	5 Q7YZF1	Q7YZF1 drosophila
29	46.5	33.2	182	5 Q7YZF0	Q7YZF0 drosophila
30	46.5	33.2	210	16 Q9RUH0	Q9RUH0 deinococcus
31	46.5	33.2	452	16 Q89PB6	Q89PB6 bradyrhizob
32	46.5	33.2	711	10 Q9FG35	Q9FG35 arabidopsis
33	46	32.9	108	5 Q8XFF6	Q8XFF6 drosophila
34	46	32.9	116	2 Q8KX54	Q8KX54 synechococc
35	46	32.9	121	17 Q8ZT42	Q8ZT42 pyrobaculum
36	46	32.9	159	16 Q8E122	Q8E122 shewanella
37	46	32.9	267	10 Q7X6S1	Q7X6S1 oryza sativ
38	46	32.9	312	16 Q88DM6	Q88DM6 pseudomonas
39	46	32.9	351	16 Q69727	Q69727 mycobacteri
40	46	32.9	351	16 Q7TVR2	Q7TVR2 mycobacteri
41	46	32.9	376	5 Q94259	Q94259 caenorhabdi
42	46	32.9	430	2 Q7WUH5	Q7WUH5 salmonella
43	46	32.9	430	16 Q8ZRN5	Q8ZRN5 salmonella
44	46	32.9	430	16 Q8Z996	Q8Z996 salmonella
45	46	32.9	540	10 Q7XYK1	Q7XYK1 chlorarachn

## ALIGNMENTS

### RESULT 1

Q94D33	PRELIMINARY;	PRT; 460 AA.
ID Q94D33		
AC Q94D33		
DT 01-DEC-2001 (TREMREL. 19, Created)		
DT 01-DEC-2001 (TREMREL. 19, Last sequence update)		
DT 01-JUN-2003 (TREMREL. 24, Last annotation update)		
DE P0712E02.23 protein (P0700A11.4 protein).		
GN P0712E02.23 OR P0700A11.4.		
OS Oryza sativa (Rice), and		
OS Oryza sativa (japonica cultivar-group).		
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;		
OC Ehrhartoideae; Oryzaceae; Oryza.		
OX NCBI_TaxID=4530, 39947;		
RN [1]		
RP SEQUENCE FROM N.A.		
RC STRAIN=cv. Nipponbare;		
RA Sasaki T., Matsumoto T., Yamamoto K.;		
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC		
RT clone:P0712E02.23;"		
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.		
RN [2]		
RP SEQUENCE FROM N.A.		
RC STRAIN=cv. Nipponbare;		
RA Sasaki T., Matsumoto T., Yamamoto K.;		
RT "Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, PAC		
RT clone:P0700A11.4;"		
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.		
EMBL; AP003492; BAB61859.1; -		
EMBL; AP003300; BAB89905.1; -		
Gramene; Q94D33; -		
GO; GO:0016020; C:membrane; IEA.		
GO; GO:0005279; F:amino acid-polyamine transporter activity; IEA.		
GO; GO:0006865; F:amino acid transport; IEA.		
InterPro; IPR002422; AA/rel_permease2.		
InterPro; IPR003006; IG_MHC.		
Pfam; PF01490; Aa trans; 1.		
PROSITE; PS00290; IG_MHC; 1.		



```

GN APT OR ALR4592.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
CC -!- FUNCTION: Catalyzes a salvage reaction resulting in the formation
CC of AMP, that is energetically less costly than de novo synthesis.
CC -!- CATALYTIC ACTIVITY: AMP + diphosphate = adenine + 5-phospho-
CC alpha-D-ribose 1-diphosphate.
CC -!- PATHWAY: Purine salvage.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the purine/pyrimidine
CC phosphoribosyltransferase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP003596; BAB76281.1; -
CC PIR; AF2378; AF2378.
CC HAMAP; MF 00004; -; 1.
CC InterPro; IPR005764; Ade_pshpso_trans.
CC InterPro; IPR002375; P2/PY_rp_transf.
CC InterPro; IPR000836; PRTtransferase.
CC Pfam; PF00156; Pribosyltran; 1.
CC TIGRFAMs; TIGR01090; apt; 1.
CC PROSITE; PS00103; PUR_PVR_PR_TRANSFER; 1.
CC Transferrase; Glycosyltransferase; Purine salvage; Complete proteome.
CC SEQUENCE 172 AA; 19144 MW; EA7762454D21AE5 CRC64;
CC -----
Query Match 32.9%; Score 46; DB 1; Length 172;
Best Local Similarity 61.1%; Pred. No. 7.9;
Matches 11; Conservative 1; Mismatches 4; Indels 2; Gaps 1;
QY 8 IVDRDPFKP--IVRSIT 23
DB 7 INDIDFPKPGILFRDIT 24
RESULT 15
ID KPT1 LVTPI STANDARD; PRT; 658 AA.
AC Q25378.
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Protein kinase C (EC 2.7.1.-).
GN PKC1.
OS Lytechinus pictus (Painted sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Echinoidea; Echinacea; Temnopleuroidea; Toxopneustidae;
OC Lytechinus.
OX NCBI_TaxID=7653;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Rakow T.L., Shen S.S.;
RT "Molecular cloning and characterization of protein kinase C from the

```

```

RT sea urchin Lytechinus pictus.";
RL Submitted (OCT-1993) to the EMBL/GenBank/DDB databases.
CC -!- FUNCTION: This is a calcium-activated, phospholipid-dependent,
CC serine- and threonine-specific enzyme (By similarity). turn
CC -!- FUNCTION: PKC is activated by diacylglycerol which in turn
CC phosphorylates a range of cellular proteins. PKC also serves as
CC the receptor for phorbol esters, a class of tumor promoters (By
CC similarity).
CC -!- SIMILARITY: Contains 2 zinc-dependent phorbol-ester and DAG
CC binding domains.
CC -!- SIMILARITY: Contains 1 C2 domain.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. PKC
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U02967; AAA03447.1; -
CC HSSP; P05697; LTEN.
CC InterPro; IPR000008; C2.
CC InterPro; IPR008973; C2_CaLB.
CC InterPro; IPR002219; DAG_PE_BIND.
CC InterPro; IPR000961; Kinase C.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR008271; Ser_Thr_pkin_AS.
CC InterPro; IPR002290; Ser_Thr_pkinase.
CC InterPro; IPR001245; Tyr_pkinase.
CC Pfam; PF00168; C2; 1.
CC Pfam; PF00130; DAG_PE_bind; 2.
CC Pfam; PF00069; kinase; 1.
CC Pfam; PF00433; kinase C; 1.
CC PRINTS; PR00360; C2DOMAIN.
CC PRINTS; PR00008; DAGPEDOMAIN.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00109; C1; 2.
CC SMART; SM00133; S_TK_X; 1.
CC SMART; SM00220; S_TK; 1.
CC PROSITE; PS00499; C2_DOMAIN_1; 1.
CC PROSITE; PS50004; C2_DOMAIN_2; 1.
CC PROSITE; PS00479; DAG_PE_BIND_DOM_1; 2.
CC PROSITE; PS00081; DAG_PE_BIND_DOM_2; 2.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC ATP-binding; Transferase; Serine/threonine-protein kinase;
CC Phorbol-ester binding; Zinc; Repeat.
CC DOMAIN 28 77 PHORBOL-ESTER AND DAG BINDING 1.
CC DOMAIN 93 142 PHORBOL-ESTER AND DAG BINDING 2.
CC DOMAIN 164 251 C2 DOMAIN.
CC DOMAIN 325 583 PROTEIN KINASE.
CC NP_BIND 331 339 ATP (BY SIMILARITY).
CC BINDING 354 354 ATP (BY SIMILARITY).
CC ACT_SITE 449 449 BY SIMILARITY.
CC SEQUENCE 658 AA; 74871 MW; 74BSA27A49C835A2 CRC64;
CC -----
Query Match 32.9%; Score 46; DB 1; Length 658;
Best Local Similarity 43.5%; Pred. No. 33;
Matches 10; Conservative 7; Mismatches 4; Indels 2; Gaps 2;
QY 5 YQSVDR-PDFPKPIVR-SITLC 25
DB 533 FQSIHEHVPSPKMSGRSVTMC 555

```

Search completed: February 26, 2004, 08:24:23  
Job time : 8.4 secs

```

or send an email to license@isb-sib.ch).
CC EMBL; AE016801; AA009792.1; -.
CC HAMAP; MF_00054; -. 1.
CC InterPro; IPR004540; EF-G.
CC InterPro; IPR000795; EF-Gbind.
CC InterPro; IPR000640; EFG C.
CC InterPro; IPR009022; EFG III V.
CC InterPro; IPR005517; EFG IV_2.
CC InterPro; IPR004161; EFTu D2.
CC InterPro; IPR005225; Small GTP.
CC InterPro; IPR009000; Translat_factor.
CC Pfam; PF00679; EFG C; 1.
CC Pfam; PF03764; EFG IV; 1.
CC Pfam; PF00009; GTP_EFTu; 1.
CC Pfam; PF03144; GTP_EFTu D2; 1.
CC PRINTS; PR00315; ELONGATNFT.
CC TIGRFAMs; TIGR00484; EF-G; 1.
CC TIGRFAMs; TIGR00231; small GTP; 1.
CC PROSITE; PS00301; EFATOR GTP; 1.
CC Elongation factor; Protein biosynthesis; GTP-binding;
KW Complete proteome.
FT NP_BIND 17 24 GTP (BY SIMILARITY).
FT NP_BIND 88 92 GTP (BY SIMILARITY).
FT NP_BIND 142 145 GTP (BY SIMILARITY).
SQ SEQUENCE 699 AA; 77413 MW; EC17B5350DDCC311 CRC64;

Query Match 33.6%; Score 47; DB 1; Length 699;
Best Local Similarity 30.0%; Pred. No. 25;
Matches 6; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGYQSVDRDPDPKPIVR 20
DB 396 CNSDKVILRMBFEPVVIQ 415

RESULT 13
ID NUOG_YERPE STANDARD; PRT; 914 AA.
AC Q8ZDL2;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE NADH-quinone oxidoreductase chain 3 (EC 1.6.99.5) (NADH dehydrogenase
I, chain G) (NDH-1, chain G).
GN NUOG OR YPO2550 OR Y1635.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]_TaxID=632;
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebatia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Felwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIMS / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matsen J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";

```

---

```

J. Bacteriol. 184:4601-4611(2002).
CC -I- FUNCTION: NDH-1 shuttles electrons from NADH, via FMN and iron-
CC sulfur (Fe-S) centers, to quinones in the respiratory chain. The
CC immediate electron acceptor for the enzyme in this species is
CC believed to be ubiquinone. Couples the redox reaction to proton
CC translocation (for every two electrons transferred, four hydrogen
CC ions are translocated across the cytoplasmic membrane), and thus
CC conserves the redox energy in a proton gradient (By similarity).
CC -I- CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.
CC -I- COFACTOR: Binds 1 2Fe-2S cluster and 3 4Fe-4S clusters per subunit
CC (By similarity).
CC -I- SUBUNIT: Composed of 13 different subunits. Subunits nuoCD, E, F,
CC and G constitute the peripheral sector of the complex.
CC -I- SIMILARITY: Belongs to the complex I 75 kDa subunit family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ414152; CAC91352.1; -.
CC EMBL; AB013767; ABM85204.1; -.
CC PIR; AD0311; AD0311.
CC InterPro; IPR009010; Asp decarb fold.
CC InterPro; IPR000283; Complex1_75K.
CC InterPro; IPR001041; Ferredoxin.
CC InterPro; IPR006963; Molybdop_Fe4S4.
CC Pfam; PF00111; fer2; 1.
CC Pfam; PF04879; Molybdop_Fe4S4; 1.
CC TIGRFAMs; TIGR01973; Nuog; 1.
CC PROSITE; PS00641; COMPLEX1_75K_1; 1.
CC PROSITE; PS00642; COMPLEX1_75K_2; 1.
CC PROSITE; PS00643; COMPLEX1_75K_3; 1.
CC Oxidoreductase; NAD; Metal-binding; Iron; Iron-sulfur; 2Fe-2S; 4Fe-4S;
KW Quinone; Ubiquinone; Complete proteome.
FT METAL 34 34 IRON-SULFUR 1 (2FE-2S) (BY SIMILARITY).
FT METAL 45 45 IRON-SULFUR 1 (2FE-2S) (BY SIMILARITY).
FT METAL 48 48 IRON-SULFUR 1 (2FE-2S) (BY SIMILARITY).
FT METAL 67 67 IRON-SULFUR 1 (2FE-2S) (BY SIMILARITY).
FT METAL 99 99 IRON-SULFUR 1 (2FE-2S) (BY SIMILARITY).
FT METAL 103 103 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT METAL 106 106 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT METAL 112 112 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT METAL 151 151 IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).
FT METAL 154 154 IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).
FT METAL 157 157 IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).
FT METAL 201 201 IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).
FT METAL 228 228 IRON-SULFUR 4 (4FE-4S) (POTENTIAL).
FT METAL 231 231 IRON-SULFUR 4 (4FE-4S) (POTENTIAL).
FT METAL 235 235 IRON-SULFUR 4 (4FE-4S) (POTENTIAL).
FT METAL 263 263 IRON-SULFUR 4 (4FE-4S) (POTENTIAL).
SQ SEQUENCE 914 AA; 100621 MW; 721487C384C2B29A CRC64;

Query Match 33.6%; Score 47; DB 1; Length 914;
Best Local Similarity 47.1%; Pred. No. 33;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 GYGQSVDRDPDPKPI 18
DB 269 GYGVLKDRPQFQOL 285

RESULT 14
APT_ANASP
ID APT_ANASP STANDARD; PRT; 172 AA.
AC Q8YN13;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Adenine phosphoribosyltransferase (EC 2.4.2.7) (APRT).

```

DZ Probable porphobilinogen deaminase (EC 2.5.1.61) (PBG)  
DE (Hydroxymethylbilane synthase) (HMBS) (Pre-uroporphyrinogen synthase).  
GN HEMC OR C1299.  
OS Chlamydia trachomatis.  
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.  
OX NCBI\_TaxID=813;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=D/UW-3/Cx;  
RX MEDLINE=9000809; PubMed=9784136;  
RA Stephens R.S., Kaiman S., Lammel C.J., Fan J., Marathe R., Aravind L.,  
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,  
RA Davis R.W.;  
RT "Genome sequence of an obligate intracellular pathogen of humans:  
RT Chlamydia trachomatis";  
RL Science 282:754-759 (1998).  
CC -!- FUNCTION: Tetrapolymerization of the monopyrrole PBG into the  
CC hydroxymethylbilane preuroporphyrinogen in several discrete steps  
CC (By similarity).  
CC -!- CATALYTIC ACTIVITY: 4 porphobilinogen + H(2)O =  
CC hydroxymethylbilane + 4 NH(3).  
CC -!- PATHWAY: Porphyrin biosynthesis by the C5 pathway; fourth step.  
CC -!- SIMILARITY: Belongs to the HMBS family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AE001302; AAC67892.1; -.  
DR PIR; E71532; E71532.  
DR HSSP; P06983; IPDA.  
DR HAMAP; MF\_00260; atypical; 1.  
DR InterPro; IPR000860; Porphobil deam.  
DR Pfam; PF01379; Porphobil deam; 1.  
DR PRINTS; PR00151; PORPHBDMNASE.  
DR ProDom; PD002745; Porphobil deam; 1.  
DR PROSITE; PS00533; PORPHOBILINOGEN DEAM; FALSE NEG.  
KW Porphyrin biosynthesis; Transferase; Complete proteome.  
SQ SEQUENCE 241 AA; 26996 MW; F89FE00C20BF7044 CRC64;  
  
Query Match 33.6%; Score 47; DB 1; Length 241;  
Best Local Similarity 47.8%; Pred. No. 7.9;  
Matches 11; Conservative 1; Mismatches 11; Indels 0; Gaps 0;  
  
QY 1 CGYGYQSIIVDRDPFKPIVRSIT 23  
Db 90 CDLGIHSKDLNPKATVWSIT 112  
  
RESULT 11  
XSC\_RHIME STANDARD; PRT; 591 AA.  
ID XSC\_RHIME  
AC Q92UW6;  
DT 15-MAR-2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Probable sulfoacetalddehyde acetyltransferase (EC 2.3.3.15).  
GN XSC OR RB0970 OR SMB21530.  
OS Rhizobium meliloti (Sinorhizobium meliloti).  
OG Plasmid pSymB (megaplasmid 2).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.  
OX NCBI\_TaxID=382;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=1021;  
RX MEDLINE=21396508; PubMed=11481431;  
RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,  
RA Vorhoefer F.J., Hernandez-Lucas I., Becker A., Gouzy J.,

RA Golding B., Puehler A.;  
RT "The complete sequence of the 1,683-kb pSymB megaplasamid from the N2-  
RT fixing endosymbiont Sinorhizobium meliloti";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894 (2001).  
CC -!- CATALYTIC ACTIVITY: Acetyl phosphate + sulfite = 2-  
CC sulfoacetalddehyde + phosphate.  
CC -!- COFACTOR: Binds 1 magnesium ion and 1 thiamine pyrophosphate per  
CC subunit (By similarity).  
CC -!- PATHWAY: Anaerobic taurine degradation; second step.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: Belongs to the TPP enzyme family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AL603645; CAC49370.1; -.  
DR PIR; B95963; B95963.  
DR InterPro; IPR000399; Pyruvate decarb.  
DR Pfam; PF00205; TPP enzymes; 1.  
DR Pfam; PF02775; TPP enzymes; C; 1.  
DR Pfam; PF02776; TPP enzymes; N; 1.  
DR PROSITE; PS00187; TPP ENZYMES; 1.  
KW Transferrase; Flavoprotein; Thiamine pyrophosphate; Metal-binding;  
KW Magnesium; Complete proteome; Plasmid.  
SQ SEQUENCE 591 AA; 64157 MW; 7234D44FB221EB97 CRC64;  
  
Query Match 33.6%; Score 47; DB 1; Length 591;  
Best Local Similarity 52.6%; Pred. No. 21;  
Matches 10; Conservative 1; Mismatches 4; Indels 4; Gaps 1;  
  
QY 1 CGYGYQSIIVDR----PDFF 15  
Db 433 CGYGPSPSIVGAKIGCPDVP 451  
  
RESULT 12  
ID EFG\_VIBVU STANDARD; PRT; 699 AA.  
AC Q8DC08;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Elongation factor G (EF-G).  
GN FUSA OR VV11338.  
OS Vibrio vulnificus.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=672;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CMCP6;  
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,  
RA Choy H.E.;  
RT "Complete genome sequence of Vibrio vulnificus CMCP6";  
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: This protein promotes the GTP-dependent translocation of  
CC the nascent protein chain from the A-site to the P-site of the  
CC ribosome.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- SIMILARITY: Belongs to the GTP-binding elongation factor family.  
CC EF-G/EF-2 subfamily.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>





```

RX MEDLINE=91378465; PubMed=1654830;
RA Carter R.S., Avadhani N.G.;
RT "Cloning and characterization of the mouse cytochrome c oxidase
  subunit IV gene.";
RL Arch. Biochem. Biophys. 288:97-106(1991).
CC -!- FUNCTION: This protein is one of the nuclear-coded polypeptide
  chains of cytochrome c oxidase, the terminal oxidase in
  mitochondrial electron transport.
CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
  c + 2 H(2)O.
CC -!- SUBCELLULAR LOCATION: Mitochondrial inner membrane.
CC -!- SIMILARITY: Belongs to the cytochrome c oxidase IV family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M37831; AAB02139.1; -.
DR EMBL; M58034; AAB02139.1; JOINED.
DR EMBL; M37829; AAB02140.1; -.
DR EMBL; X54591; CAA38507.1; -.
DR PIR; S12142; S12142.
DR MGD; MGI:88473; Cox41l.
DR GO; GO:0005739; C:mitochondrion; IDA.
DR InterPro; IPR004203; COX4.
DR Pfam; PF02936; COX4; 1.
KW Oxidoreductase; Inner membrane; Mitochondrion; Transist peptide.
FT TRANSIT 1 22 CYTOCHROME C OXIDASE SUBUNIT IV ISOFORM
FT CHAIN 23 169
FT CONFLICT 34 34 A -> R (IN REF. 1).
FT CONFLICT 71 71 S -> N (IN REF. 1).
FT CONFLICT 140 141 MQ -> IE (IN REF. 1; AAB02139 AND 2).
SQ SEQUENCE 169 AA; 19530 MW; D30BLDBE14FDBEA CRC64;

Query Match 34.3%; Score 48; DB 1; Length 169;
Best Local Similarity 36.4%; Pred. No. 3.8;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 3 YGYQSVDRPDPFKPIVRSITL 24
DB 33 YAFPTYADRDYPLDVAHVM 54

RESULT 7
PTA_THEME
ID_PTA_THEME STANDARD; PRT; 294 AA.
AC Q9X0L4;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-MAY-2004 (Rel. 43, Last annotation update)
DE Phosphate acetyltransferase (EC 2.3.1.8) (Phosphotransacetylase).
GN PTA OR TM1130.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]_TaxID=2336;
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109 / ATCC 43589;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
  Haft D.H., Hickey E.K., Peterson J.D., Linher K.D., Garrett M.M.,
  McDonald L., Utterback T.R., Malek J.A., Phillips C.A., Richardson D.,
  Stewart A.M., Sutton M.D., Pratt M.S., Heidelberg J., Sutton G.G.,
  Fleischmann R.D., Eisen J.A., White O., Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
  "Evidence for lateral gene transfer between Archaea and Bacteria from
  genome sequence of Thermotoga maritima.";
  Nature 399:323-329(1999).

```

```

RN SEQUENCE OF 1-39, AND CHARACTERIZATION.
RP MEDLINE=99173910; PubMed=10074080;
RA Beck A.-K., Glasmecher J., Schmidt R., Schoenheit P.;
RT "Purification and characterization of two extremely thermostable
  enzymes, phosphate acetyltransferase and acetate kinase, from the
  hyperthermophilic eubacterium Thermotoga maritima.";
RL J. Bacteriol. 181:1861-1867(1999).
CC -!- FUNCTION: IN ADDITION TO ACETYL-COA (100%), THE ENZYME ACCEPTS
  PROPIONYL-COA (60%) AND BUTYRYL-COA (30%). HAS A TEMPERATURE
  OPTIMUM AT 90 DEGREES CELSIUS.
CC -!- CATALYTIC ACTIVITY: Acetyl-CoA + phosphate = CoA + acetyl
  phosphate.
CC -!- PATHWAY: Conversion of acetate to acetyl-CoA; second step.
CC -!- SUBUNIT: Homotetramer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- SIMILARITY: BELONGS TO THE PHOSPHATE ACETYLTRANSFERASE AND
  BUTYRYLTRANSFERASE FAMILY. MORE SIMILAR TO PTB THAN PTA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; A8001770; AAD36206.1; -.
DR PIR; G72293; G72293.
DR TIGR; TM1130; -.
DR InterPro; IPR002505; PTA_PTB.
DR Pfam; PF01515; PTA_PTB; 1.
KW Transferase; Acyltransferase; Complete proteome.
FT CONFLICT 11 11 I -> Y (IN REF. 2).
SQ SEQUENCE 294 AA; 32093 MW; F4B98B3CAE120AFB CRC64;

Query Match 34.3%; Score 48; DB 1; Length 294;
Best Local Similarity 46.7%; Pred. No. 6.9;
Matches 7; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 7 SIVDRPDPFKPIVRS 21
DB 122 SVMETPDPFRPLIIS 136

RESULT 8
EFG_VIBPA
ID_EFG_VIBPA STANDARD; PRT; 699 AA.
AC Q87L45;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Elongation factor G (EF-G).
GN FUSA OR VP2771.
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
  Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]_TaxID=670;
RP SEQUENCE FROM N.A.
RC STRAIN=RIMD 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
  Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Khmura S.,
  Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
  "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
  distinct from that of V. cholerae.";
  Lancet 361:743-749(2003).
CC -!- FUNCTION: This protein promotes the GTP-dependent translocation of
  the nascent protein chain from the A-site to the P-site of the
  ribosome.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the GTP-binding elongation factor family.

```

```
SQ SEQUENCE 705 AA; 77866 MW; 8D75AA4BEB102499 CRC64;
Query Match 35.0%; Score 49; DB 1; Length 705;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 5 YQSVIVDRPDPFPKPIVRSI 22
   :|||:|:|:|
DB 513 YNEVDRPVKPTPVVAGI 530

RESULT 4
MK32 YEAST
ID MK32 YEAST STANDARD; PRT; 363 AA.
AC P23060;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE MAK32 protein.
GN MAK32 OR YCR019W OR YCR19W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89131254; PubMed=3916862;
RA Toh-E A., Sahashi Y.;
RT "The PRT18 locus of Saccharomyces cerevisiae: a complex locus
   containing multiple genes.";
RL Yeast 1:159-171(1985).
RN [2]
RP SEQUENCE FROM N.A.
RA Feldmann H., Mannhaupt G., Vetter I.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Necessary for the structural stability of L-A double-
   stranded RNA-containing particles. Necessary for growth at 37
   degrees Celsius as well as for maintenance of the killer plasmid.
CC -!- SIMILARITY: TO S.POMBE SPAC468.14C.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL outstation -
   the European Bioinformatics Institute. There are no restrictions on its
   use by non-profit institutions as long as its content is in no way
   modified and this statement is not removed. Usage by and for commercial
   entities requires a license agreement (See http://www.isb-sib.ch/announce/
   or send an email to license@isb-sib.ch).
CC
DR EMBL; X59720; CRA42310.1; -.
DR PIR; S19429; BVBYK2.
DR GenOnline; 138924; -.
DR SGD; S0000612; MAK32.
DR GO; GO:0019048; P:virus-host interaction; IMP.
FT CONFLICT 15 15 I->II (IN REF. 1).
FT CONFLICT 82 82 MISSING (IN REF. 1).
SQ SEQUENCE 363 AA; 40783 MW; CFE358F8ACF6EAAC CRC64;
Query Match 34.6%; Score 48.5; DB 1; Length 363;
Best Local Similarity 55.0%; Pred. No. 7.2;
Matches 11; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

QY 4 GYQSVIVR-PDPFPKPIVRSI 22
   :|||:|:|:|
DB 61 GLKWIVDGRDPPKPEVIREI 80

RESULT 5
CX41 PITPI
ID CX41 PITPI STANDARD; PRT; 144 AA.
AC O46585;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
```

```
DE DE Cytochrome c oxidase subunit IV isoform 1 (EC 1.9.3.1) (COX IV-1)
DE COX41 OR COX4.
OS Pithecia pithecia (White-faced saki).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Pitheciinae;
OC Pithecia.
OX NCBI_TaxID=43777;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97277139; PubMed=9115172;
RA Wu W., Goodman M., Lonax M.I., Grossman L.I.;
RT "Molecular evolution of cytochrome c oxidase subunit IV: evidence for
   positive selection in simian primates.";
RL J. Mol. Evol. 44:477-491(1997).
CC -!- FUNCTION: This protein is one of the nuclear-coded polypeptide
   chains of cytochrome c oxidase, the terminal oxidase in
   mitochondrial electron transport.
CC -!- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
   c + 2 H(2)O.
CC -!- SUBCELLULAR LOCATION: Mitochondrial inner membrane.
CC -!- SIMILARITY: Belongs to the cytochrome c oxidase IV family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
   between the Swiss Institute of Bioinformatics and the EMBL outstation -
   the European Bioinformatics Institute. There are no restrictions on its
   use by non-profit institutions as long as its content is in no way
   modified and this statement is not removed. Usage by and for commercial
   entities requires a license agreement (See http://www.isb-sib.ch/announce/
   or send an email to license@isb-sib.ch).
CC
DR EMBL; AF042770; AAB97849.1; -.
DR EMBL; AF042768; AAB97849.1; JOINED.
DR EMBL; AF042769; AAB97849.1; JOINED.
DR InterPro; IPR004203; COX4.
DR Pfam; PF02936; COX4; 1.
KW Oxidoreductase; Inner membrane; Mitochondrion.
FT NON_TER 1
FT SEQUENCE 144 AA; 16917 MW; 5573133B773B5C89 CRC64;
SQ SEQUENCE 144 AA; 16917 MW; 5573133B773B5C89 CRC64;
Query Match 34.3%; Score 48; DB 1; Length 144;
Best Local Similarity 50.0%; Pred. No. 3.2;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 3 YGVQSVIVDRPDPFPKPIVRSI 22
   :|||:|:|:|
DB 8 YTLQSVIVDRDPLPDVAHV 27

RESULT 6
CX41 MOUSE
ID CX41 MOUSE STANDARD; PRT; 169 AA.
AC P19783;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Cytochrome c oxidase subunit IV isoform 1, mitochondrial precursor
   (EC 1.9.3.1) (COX IV-1) (Cytochrome c oxidase polypeptide IV).
DE COX41 OR COX4A OR COX4.
GN Mus musculus (Mouse).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BALB/c; TISSUE=Heart;
RA Grossman L.I., Akamatsu M.;
RT "Nucleotide sequence of a mouse cDNA for subunit IV of cytochrome c
   oxidase.";
RL Nucleic Acids Res. 18:6454-6454(1990).
RN [2]
RP SEQUENCE FROM N.A.
```

```

FT DOMAIN 305 421 CH4.
FT DISULFID 23 75 BY SIMILARITY.
FT DISULFID 121 180 BY SIMILARITY.
FT DISULFID 226 285 BY SIMILARITY.
FT DISULFID 330 392 BY SIMILARITY.
FT CARBOHYD 43 43 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 95 95 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 261 261 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 365 365 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 415 415 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 421 AA; 47320 MW; 8F909E1F30A06B47 CRC64;

Query Match 77.1%; Score 108; DB 1; Length 421;
Best Local Similarity 90.9%; Pred. No. 6.2e-09;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GYGYSIVDRPDPFKPIVRSIT 23
| | | | | | | | | | | | | | | | | | | | |
DB 280 GYGYSIVDRPDPFKPIVRSIT 301

RESULT 2
EPC_RAT
ID EPC_RAT STANDARD; PRT; 429 AA.
AC P01855;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ig epsilon chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (IMMUNOCYTOMA IR2).
RC STRAIN=LOU/C/WSL;
RX MEDLINE=83064537; PubMed=6292865;
RA Hellman L., Pettersson U., Engstroem A., Karlsson T., Bennich H.;
RT "Structure and evolution of the heavy chain from rat immunoglobulin
RT E.";
RL Nucleic Acids Res. 10:6041-6049(1982).
RN [2]
RP SEQUENCE OF 168-342 FROM N.A. (MYELOMA IR162).
RX MEDLINE=83182019; PubMed=6820340;
RA Kindsvogel W.R., Reddy E.P., Moore J.M., Faust C.H. Jr.;
RT "A cloned cDNA probe for rat immunoglobulin epsilon heavy chain:
RT construction, identification, and DNA sequence.";
RL DNA 1:335-343(1982).
RN [3]
RP SEQUENCE OF 205-306 FROM N.A.
RX MEDLINE=82174576; PubMed=6803238;
RA Hellman L., Pettersson U., Bennich H.;
RT "Characterization and molecular cloning of the mRNA for the heavy
RT (epsilon) chain of rat immunoglobulin E.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:1264-1268(1982).
CC -1- SIMILARITY: Contains 4 immunoglobulin-like domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J00744; AAA41379.1; ALT_INIT.
DR PIR; A93442; EHRT.
DR HSSP; P01854; IIGE.
DR InterPro; IPR007110; Ig-like.

```

```

DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; ig; 4.
DR SMART; SM00407; Igcl; 1.
DR PROSITE; PS00835; IG_LIKE; 4.
DR PROSITE; PS00290; IG_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region; Repeat.
FT NON_TER 1 1
FT DOMAIN 6 89 IG-LIKE 1.
FT DOMAIN 103 201 IG-LIKE 2.
FT DOMAIN 205 305 IG-LIKE 3.
FT DOMAIN 314 414 IG-LIKE 4.
FT CONFLICT 168 168 R -> N (IN REF. 2).
FT CONFLICT 308 308 P -> L (IN REF. 2).
SQ SEQUENCE 429 AA; 48671 MW; D2970B34EF8A72B0 CRC64;

Query Match 60.7%; Score 85; DB 1; Length 429;
Best Local Similarity 77.3%; Pred. No. 2.2e-05;
Matches 17; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 GYGYSIVDRPDPFKPIVRSIT 23
| | | | | | | | | | | | | | | | | | | | |
DB 284 GEGYQCRVDHPHPFKPIVRSIT 305

RESULT 3
PURL_PYRAB
ID PURL_PYRAB STANDARD; PRT; 705 AA.
AC Q9UXW6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Phosphoribosylformylglycinamide synthase II (EC 6.3.5.3) (FGAM
DE synthase II).
GN PURL OR PYRAB17420 OR PAB1201.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=29292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GES / Orsay;
RX MEDLINE=22511545; PubMed=12622808;
RA Cohen G.N., Barbe V., Flament D., Galperin M., Heilig R., Lecompte O.,
RA Poch O., Prieur D., Querellou J., Ripp R., Thierry J.-C.,
RA Van der Oost J., Weissenbach J., Zivanovic Y., Forterre P.;
RT "An integrated analysis of the genome of the hyperthermophilic
RT archaeon Pyrococcus abyssi.";
RL Microbiol. 47:1495-1512(2003).
CC -1- CATALYTIC ACTIVITY: ATP + N(2)-(5-phospho-D-ribose) + 2-
CC ribosylglycinamide + L-glutamine + H(2)O = ADP + phosphate + 2-
CC (formamido)-N(1)-(5-phospho-D-ribose)acetamide + L-glutamate.
CC
CC -1- PATHWAY: De novo purine biosynthesis; fourth step.
CC -1- SUBUNIT: Heterodimer of two subunits; purQ and purL.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the FGAMS family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ248288; CAB50647.1; -.
DR PIR; A75026; A75026.
DR HAMAP; MF_00420; -.
DR InterPro; IPR000728; AIR_synth.
DR Pfam; PF00586; AIRS; 2.
DR Pfam; PF02769; AIRS_C; 2.
KW Purine biosynthesis; Ligase; ATP-binding; Complete proteome.
NP_BIND 89 100

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 08:17:40 ; Search time 6.4 Seconds  
(without alignments)  
203.399 Million cell updates/sec

Title: US-09-701-623c-8  
Perfect score: 140  
Sequence: 1 CGYGYQIVDRPFPKPIVRSITLC 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 segs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	108	77.1	421	1	EPC MOUSE
2	85	60.7	429	1	EPC RAT
3	49	35.0	705	1	PURL PYRAB
4	48.5	34.6	363	1	MK32 YEAST
5	48	34.3	144	1	OX41 PITPI
6	48	34.3	169	1	OX41 MOUSE
7	48	34.3	294	1	PTA THEMA
8	48	34.3	699	1	EFG VIBPA
9	47	33.6	169	1	OX41 RAT
10	47	33.6	241	1	HEM3 CHLTR
11	47	33.6	591	1	XSC RHIME
12	47	33.6	699	1	EFG VIBU
13	47	33.6	914	1	NUOG VERPE
14	46	32.9	172	1	APT ANASP
15	46	32.9	658	1	KPCI LYTPI
16	46	32.9	705	1	PURL PYRHO
17	46	32.9	909	1	NUOG SHEON
18	45	32.1	172	1	APT SYNY3
19	45	32.1	273	1	TRME SCHPO
20	45	32.1	294	1	ECMC MYCLE
21	45	32.1	698	1	EFGI SHEON
22	45	32.1	907	1	NUOG ECO57
23	45	32.1	907	1	NUOG ECOL6
24	45	32.1	907	1	NUOG ECOLI
25	45	32.1	907	1	NUOG ECOLI
26	45	32.1	907	1	NUOG SHIFL
27	45	32.1	925	1	NFC2 HUMAN
28	44.5	31.8	62	1	STAT HUMAN
29	44.5	31.8	222	1	AMYB SECC
30	44.5	31.8	335	1	Y094 CAEEL
31	44.5	31.8	1056	1	YNN2 YEAST
32	44	31.4	55	1	OX41 CEBAP
33	44	31.4	137	1	OX41 PERPO

RESULT 1

ID	EPC_MOUSE	STANDARD;	PRT;	421 AA.
AC	P06336; P01856;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	IG epsilon chain C region.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=84236092; PubMed=6329729;			
RA	Ishida N., Ueda S., Hayashida H., Miyata T., Honjo T.;			
RT	"The nucleotide sequence of the mouse immunoglobulin epsilon gene:			
RT	comparison with the human epsilon gene sequence.";			
RL	EMBO J. 1:1117-1123 (1982).			
RN	[2]			
RP	REVISIONS.			
RA	Honjo T.;			
RL	Submitted (APR-1986) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE OF 34-421 FROM N.A.			
RX	MEDLINE=83117774; PubMed=6818553;			
RA	Liu F.-T., Albrandt K., Sutcliffe J.G., Katz D.H.;			
RT	"Cloning and nucleotide sequence of mouse immunoglobulin epsilon			
RT	chain cDNA.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 79:7852-7856 (1982).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; X01857; CAA25977.1; -			
DR	EMBL; X01857; CAA25978.1; -			
DR	PIR; A02144; EHM5			
DR	PIR; A02145; EHM5S			
DR	HSP; P01854; IIG.			
DR	InterPro; IPR007110; Ig-like.			
DR	InterPro; IPR003597; Ig_c1.			
DR	InterPro; IPR003006; Ig_MHC.			
DR	Pfam; PF00047; Ig; 4.			
DR	SMART; SM00407; IGc1; 2.			
DR	PROSITE; PS00835; IG_LIKE; 4.			
DR	PROSITE; PS00290; IG_MHC; 3.			
KW	Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.			
FT	NON TER 1 90			CH1.
FT	DOMAIN 1 197			CH2.
FT	DOMAIN 91 304			CH3.

O46584 aotus azara  
P00423 bos taurus  
Q9ttt8 oryctolagus  
Q8dgt9 synechococ  
P78795 schizosacch  
Q8kcd6 chlorobium  
Q8kts9 candidatus  
P14133 cucumis sat  
Q9hwd2 pseudomonas  
Q88rn8 pseudomonas  
P55849 mus musculu  
P54679 dictyosteli

## ALIGNMENTS

A.Accession: S14190  
A.Status: nucleic acid sequence not shown; translation not shown  
A.Molecule type: mRNA  
A.Residues: 1-169 <VR>  
A.Cross-references: AID:X54081; NID:G57030; PIDN:CAA38018.1; PID:G57031  
A.Experimental source: strain Sprague Dawley  
A.Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1990  
R.Schaeffer, H.; Noack, H.; Halingk, W.; Brandt, U.; von Jagow, G.  
Eur. J. Biochem. 230, 235-241, 1995  
A.Title: Cytochrome-c oxidase in developing rat heart. Enzymic properties and amino-term  
A.Reference number: S65372; MUID:95324529; PMID:7601105  
A.Accession: S65373  
A.Status: preliminary  
A.Molecule type: protein  
A.Residues: 23-45 <SCH>  
C.Genetics:  
A.Gene: RCO4-1  
A.Introns: 25/1; 81/1; 125/1  
C.Function:  
A.Description: the cytochrome-c oxidase complex catalyzes the oxidation of four molecules  
ns from the mitochondrial matrix producing two molecules of water and lowering the conce  
A.Pathway: oxidative phosphorylation; respiratory chain  
C.Superfamily: cytochrome-c oxidase chain IV  
C.Keywords: electron transfer; membrane-associated complex; mitochondrial inner membrane  
F1-22/Domin: transit peptide (mitochondrion) #status predicted <NP>  
F123-169/Product: cytochrome-c oxidase chain IV #status experimental <MAT>  
F177-103/Domin: transmembrane helix #status predicted <TR01>

```
Query Match      33.6%; Score 47; DB 1; Length 169;
Best Local Similarity 45.5%; Pred. No. 10;
Matches 10; Conservative 2; Mismatches 10; Indels 0; Gaps 0;
```

QY 3 YGYQSVDRDPFKPIVRSITL 24  
DB 33 YALPSYVDRRDYPLEDVAHVKL 54

Search completed: February 26, 2004, 08:28:16  
Job time : 12. secs

T06521  
pitrilysin (EC 3.4.24.55) - garden pea  
C:Species: Pisum sativum (garden pea)  
C>Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
C:Accession: T06521  
R;Vandervere, P.S.; Bennett, T.M.; Oblong, J.E.; Lamppa, G.K.  
Proc. Natl. Acad. Sci. U.S.A. 92, 7177-7181, 1995  
A:Title: A chloroplast processing enzyme involved in precursor maturation shares a zinc.  
A:Reference number: Z15733; MUID:95365331; PMID:7638164  
A:Accession: T06521  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1259 <VAN>  
A:Cross-references: EMBL:U25111; NID:g1065907; PIDN:AAA81472.1; PID:g799369  
C:Genetics:  
A:Genome: nuclear  
C:Function:  
A:Description: catalyzes proteolytic removal of chloroplast transit peptides  
C:Keywords: hydrolase; metalloproteinase; zinc

Query Match 34.3%; Score 48; DB 2; Length 1259;  
Best Local Similarity 38.1%; Pred. No. 68;  
Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 5 YOSIVDRPDPKPIVRSITLC 25  
DB 79 YSVLSQPTAPVPRQSTSC 99

RESULT 15  
A3209  
cytochrome-c oxidase (EC 1.9.3.1) chain IV precursor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 10-Aug-1990 #sequence\_revision 10-Aug-1990 #text\_change 11-Jun-1999  
C:Accession: S12724; S04070; S04593; S14190; S65373  
R;Yamada, M.; Amuro, N.; Goto, Y.; Okazaki, T.  
J. Biol. Chem. 265, 7687-7692, 1990  
A:Title: Structural organization of the rat cytochrome c oxidase subunit IV gene.  
A:Reference number: A35209; MUID:90237079; PMID:2159010  
A:Accession: A35209  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-169 <YAM>  
A:Cross-references: GB:J05425; NID:g203516; PIDN:AAA40949.1; PID:g203517  
R;Amuro, N.; Yamada, M.; Goto, Y.; Okazaki, T.  
Nucleic Acids Res. 18, 3992, 1990  
A:Title: Complete nucleotide sequence of the gene encoding rat cytochrome c oxidase subunit IV.  
A:Reference number: S12724; MUID:90236528; PMID:2165254  
A:Accession: S12724  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-169 <AMU>  
A:Cross-references: EMBL:J05425; NID:g203516; PIDN:AAA40949.1; PID:g203517  
R;Goto, Y.; Amuro, N.; Okazaki, T.  
Nucleic Acids Res. 17, 2851, 1989  
A:Title: Nucleotide sequence of cDNA for rat brain and liver cytochrome c oxidase subunit IV.  
A:Reference number: S04070; MUID:89240039; PMID:2541414  
A:Accession: S04070  
A:Molecule type: mRNA  
A:Residues: 1-169 <GOT>  
A:Cross-references: EMBL:X14209; NID:g55989; PIDN:CAA32426.1; PID:g55990  
R;Gopalan, G.; Droste, M.; Kadenbach, B.  
Nucleic Acids Res. 17, 4376, 1989  
A:Title: Nucleotide sequence of cDNA encoding subunit IV of cytochrome c oxidase from f.  
A:Reference number: S04593; MUID:89296488; PMID:2544859  
A:Accession: S04593  
A:Molecule type: mRNA  
A:Residues: 1-169 <GOP>  
A:Cross-references: EMBL:X15029; NID:g55980; PIDN:CAA33133.1; PID:g55981  
R;Virbasius, J.V.; Scarpulla, R.C.  
Nucleic Acids Res. 18, 6581-6586, 1990  
A:Title: The rat cytochrome c oxidase subunit IV gene family: tissue-specific and hormonal regulation.  
A:Reference number: S14190; MUID:91067442; PMID:2174541

Query Match 34.3%; Score 48; DB 1; Length 169;  
Best Local Similarity 36.4%; Pred. No. 7.2;  
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 3 YXQSVDRPDPKPIVRSITL 24  
DB 33 YAFPIYADRRDPLPDVAHVTM 54

RESULT 13  
G72293  
phosphate butyryltransferase - Thermotoga maritima (strain MSB8)  
C:Species: Thermotoga maritima  
C>Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
C:Accession: G72293  
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Dodson, R.J.; Haft, D.H.; Hickey, J.C.; Garret, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.  
Nature 399, 323-329, 1999  
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing of *Thermotoga maritima*.  
A:Reference number: A72200; MUID:99287316; PMID:10360571  
A:Accession: G72293  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-294 <ARN>  
A:Cross-references: GB:AE001770; GB:AE000512; NID:g4981658; PIDN:AAD36206.1; PID:g498167  
A:Experimental source: strain MSB8  
C:Genetics:  
A:Gene: TM130  
C:Superfamily: phosphate acetyltransferase

Query Match 34.3%; Score 48; DB 2; Length 294;  
Best Local Similarity 46.7%; Pred. No. 13;  
Matches 7; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 7 SIVDRPDPKPIVRS 21  
DB 122 SYMEIPDRPRLIIS 136

RESULT 14

T06521  
pitrilysin (EC 3.4.24.55) - garden pea  
C:Species: Pisum sativum (garden pea)  
C>Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
C:Accession: T06521  
R;Vandervere, P.S.; Bennett, T.M.; Oblong, J.E.; Lamppa, G.K.  
Proc. Natl. Acad. Sci. U.S.A. 92, 7177-7181, 1995  
A:Title: A chloroplast processing enzyme involved in precursor maturation shares a zinc.  
A:Reference number: Z15733; MUID:95365331; PMID:7638164  
A:Accession: T06521  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1259 <VAN>  
A:Cross-references: EMBL:U25111; NID:g1065907; PIDN:AAA81472.1; PID:g799369  
C:Genetics:  
A:Genome: nuclear  
C:Function:  
A:Description: catalyzes proteolytic removal of chloroplast transit peptides  
C:Keywords: hydrolase; metalloproteinase; zinc

Query Match 34.3%; Score 48; DB 2; Length 1259;  
Best Local Similarity 38.1%; Pred. No. 68;  
Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 5 YOSIVDRPDPKPIVRSITLC 25  
DB 79 YSVLSQPTAPVPRQSTSC 99

RESULT 15  
A3209  
cytochrome-c oxidase (EC 1.9.3.1) chain IV precursor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 10-Aug-1990 #sequence\_revision 10-Aug-1990 #text\_change 11-Jun-1999  
C:Accession: S12724; S04070; S04593; S14190; S65373  
R;Yamada, M.; Amuro, N.; Goto, Y.; Okazaki, T.  
J. Biol. Chem. 265, 7687-7692, 1990  
A:Title: Structural organization of the rat cytochrome c oxidase subunit IV gene.  
A:Reference number: A35209; MUID:90237079; PMID:2159010  
A:Accession: A35209  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-169 <YAM>  
A:Cross-references: GB:J05425; NID:g203516; PIDN:AAA40949.1; PID:g203517  
R;Amuro, N.; Yamada, M.; Goto, Y.; Okazaki, T.  
Nucleic Acids Res. 18, 3992, 1990  
A:Title: Complete nucleotide sequence of the gene encoding rat cytochrome c oxidase subunit IV.  
A:Reference number: S12724; MUID:90236528; PMID:2165254  
A:Accession: S12724  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-169 <AMU>  
A:Cross-references: EMBL:J05425; NID:g203516; PIDN:AAA40949.1; PID:g203517  
R;Goto, Y.; Amuro, N.; Okazaki, T.  
Nucleic Acids Res. 17, 2851, 1989  
A:Title: Nucleotide sequence of cDNA for rat brain and liver cytochrome c oxidase subunit IV.  
A:Reference number: S04070; MUID:89240039; PMID:2541414  
A:Accession: S04070  
A:Molecule type: mRNA  
A:Residues: 1-169 <GOT>  
A:Cross-references: EMBL:X14209; NID:g55989; PIDN:CAA32426.1; PID:g55990  
R;Gopalan, G.; Droste, M.; Kadenbach, B.  
Nucleic Acids Res. 17, 4376, 1989  
A:Title: Nucleotide sequence of cDNA encoding subunit IV of cytochrome c oxidase from f.  
A:Reference number: S04593; MUID:89296488; PMID:2544859  
A:Accession: S04593  
A:Molecule type: mRNA  
A:Residues: 1-169 <GOP>  
A:Cross-references: EMBL:X15029; NID:g55980; PIDN:CAA33133.1; PID:g55981  
R;Virbasius, J.V.; Scarpulla, R.C.  
Nucleic Acids Res. 18, 6581-6586, 1990  
A:Title: The rat cytochrome c oxidase subunit IV gene family: tissue-specific and hormonal regulation.  
A:Reference number: S14190; MUID:91067442; PMID:2174541

## Extremophiles 2, 417-425, 1998

A:Title: Genetic profile of pNOB8 from Sulfolobus: the first conjugative plasmid from an  
 A:Reference number: Z20959; MUID:99044580; PMID:19827331  
 A:Accession: T31043  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-165 <SHE>  
 A:Cross-references: EMBL:AJ010405; NID:e1351926; PID:e1351965; PIDN:CAA09149.1  
 A:Experimental source: strain NOB8H2  
 C:Genetics:  
 A:Genome: plasmid pNOB8

Query Match 35.0%; Score 49; DB 2; Length 165;

Best Local Similarity 58.8%; Pred. No. 4.9;

Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 5 YQIVDRDPDFPKPIVRS 21

Db 71 YERIDKDFNPRPIVRS 87

## RESULT 8

AG2525  
 Hypothetical protein alr7383 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120a1  
 C:Species: Nostoc sp. PCC 7120  
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
 A:Accession: AG2525  
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watarabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
 DNA Res. 8, 205-213, 2001  
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena  
 A:Reference number: AB1807; MUID:21595285; PMID:11759840  
 A:Accession: AG2525  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-232 <KUP>  
 A:Cross-references: GB:BA000020; PIDN:BAB77141.1; PID:gl7134582; GSPDB:GN00180  
 A:Experimental source: strain PCC 7120  
 C:Genetics:  
 A:Gene: alr7383  
 A:Genome: plasmid

Query Match 35.0%; Score 49; DB 2; Length 232;

Best Local Similarity 43.8%; Pred. No. 7.2;

Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 5 YQIVDRDPDFPKPIVR 20

Db 49 YAAVVDHPDYNVCVK 64

## RESULT 9

C75354  
 conserved hypothetical protein - Deinococcus radiodurans (strain R1)  
 C:Species: Deinococcus radiodurans  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
 A:Accession: C75354  
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A:Reference number: A75250; MUID:20036896; PMID:10567256  
 A:Accession: C75354  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-433 <WHI>  
 A:Cross-references: GB:AE0002019; GB:AE000513; NID:g6459547; PIDN:AAE11328.1; PID:g645954  
 A:Experimental source: strain R1  
 C:Genetics:  
 A:Gene: DR1773  
 A:Map position: 1

Query Match 35.0%; Score 49; DB 2; Length 433;

Best Local Similarity 35.3%; Pred. No. 14;

Matches 6; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

Query Match 35.0%; Score 49; DB 2; Length 433;

Best Local Similarity 35.3%; Pred. No. 14;

Matches 6; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

Qy 6 QSVIVDRDPDFPKPIVRSI 22

Db 117 EAVAQRPDLPPELLRQL 133

## RESULT 10

A75026  
 phosphoribosylformylglycinamide synthase II (purl) PAB1201 - Pyrococcus abyssi (strain C:Species: Pyrococcus abyssi  
 C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
 A:Accession: A75026  
 R:anonymous, Genoscope  
 submitted to the EMBL Data Library, July 1999  
 A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome str  
 A:Reference number: A75001  
 A:Accession: A75026  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-705 <XAW>  
 A:Cross-references: GB:AJ248288; GB:AL096836; NID:g5458960; PIDN:CAB50647.1; PID:g54591  
 A:Experimental source: strain Orsay  
 C:Genetics:  
 A:Gene: PAB1201  
 C:Superfamily: phosphoribosylformylglycinamide synthase component II

Query Match 35.0%; Score 49; DB 2; Length 705;

Best Local Similarity 50.0%; Pred. No. 25;

Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 5 YQIVDRDPDFPKPIVRSI 22

Db 513 YNEVDRPVKFTPVVAGI 530

## RESULT 11

BVBYK2  
 MAK32 protein - Yeast (Saccharomyces cerevisiae)  
 N:Alternate names: protein YCR019w  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 30-Jun-1991 #sequence\_revision 30-Sep-1993 #text\_change 16-Jun-2000  
 A:Accession: S19429; S07695  
 R:Feldmann, H.; Manhaupt, G.; Vetter, I.  
 submitted to the Protein Sequence Database, March 1992  
 A:Reference number: S19429  
 A:Accession: S19429  
 A:Molecule type: DNA  
 A:Residues: 1-363 <FEL>  
 A:Cross-references: EMBL:X59720; NID:gl907116; PIDN:CAA42310.1; PID:gl907162; GSPDB:GNO  
 R:Toh-e, A.; Sahashi, Y.  
 Yeast 1, 159-171, 1985  
 A:Title: The PET18 locus of Saccharomyces cerevisiae: a complex locus containing multip  
 A:Reference number: S07692; MUID:89131254; PMID:3916862  
 A:Accession: S07695  
 A:Molecule type: DNA  
 A:Residues: 1-14,15-81,83-282,18',284-363 <TOH>  
 C:Genetics:  
 A:Gene: SGD:MAK32; MIPS:YCR019w  
 A:Cross-references: SGD:S0000612; MIPS:YCR019w  
 A:Map position: 3R  
 C:Superfamily: MAK32 protein

Query Match 34.6%; Score 48.5; DB 1; Length 363;

Best Local Similarity 55.0%; Pred. No. 14;

Matches 11; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

Qy 4 GYQIVDR-PDRPKPIVRSI 22

Db 61 GLKWIVDRGSDPFKEVIREI 80

```

Qy      2  GYGQSI VDRPDPFKPI VRSIT  23
      ||||| ||| ||||| |||||
Db      247 GYGQCI VDRPDPFKPI VRSIT  268

RESULT  5
S38864

```

RESULT 7  
T31043  
hypothetical protein 165 - *Sulfolobus* sp. plasmid pNOB8  
C;Species: *Sulfolobus* sp.  
C;Date: 02-Sep-2000 #sequence\_revision 02-Sep-2000 #text\_change 02-Sep-2000  
C;Accession: T31043  
R;She. O.; Phan. H.; Garrett, R.A.; Albers, S.V.; Stedman, K.M.; Zillig, W.



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 08:18:56 ; Search time 10 seconds  
(without alignments)  
240.479 Million cell updates/sec

Title: US-09-701-623C-8  
Perfect score: 140  
Sequence: 1 GYGQSVIVRPPDPKPIVRSITL 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	116	82.9	423	1 EHMS	Ig epsilon chain C
2	108	77.1	107	2 I68730	IgE chain C3 regio
3	108	77.1	107	2 I68726	IgE chain C3 regio
4	108	77.1	388	1 EHMS	Ig epsilon chain C
5	108	77.1	548	2 S38864	Ig epsilon chain C
6	85	60.7	429	1 EHRT	Ig epsilon chain C
7	49	35.0	165	2 T31043	hypothetical prote
8	49	35.0	232	2 AG2525	hypothetical prote
9	49	35.0	433	2 C73354	conserved hypothet
10	49	35.0	705	2 A75026	phosphoribosylform
11	48.5	34.6	363	1 BVBXK2	MAX32 protein - ye
12	48	34.3	169	1 S12142	cytochrome-c oxida
13	48	34.3	294	2 G72293	phosphate butyrylt
14	48	34.3	1259	2 T05521	pirilysin (EC 3.4
15	47	33.6	169	1 A35209	cytochrome-c oxida
16	47	33.6	241	2 E71532	probable porphobil
17	47	33.6	591	2 B95963	probable acetolact
18	47	33.6	914	2 AD0311	NADH2 dehydrogenas
19	46.5	33.2	210	2 C75399	hypothetical prote
20	46	32.9	172	2 AF2378	adenine phosphorib
21	46	32.9	351	2 P70801	probable fad336 pr
22	46	32.9	359	2 P89075	protein R04A8.1 [i
23	46	32.9	426	2 I36948	Ig epsilon-chain -
24	46	32.9	430	2 AG0531	cell cycle protein
25	46	32.9	705	2 A71211	probable phosphori
26	46	32.9	715	2 S54628	hypothetical prote
27	46	32.1	172	2 S75440	adenine phosphorib
28	45	32.1	294	1 S72706	probable enoyl-CoA
29	45	32.1	910	2 G91024	NADH dehydrogenase

NADH dehydrogenase  
NADH2 dehydrogenas  
NADH2 dehydrogenas  
transcription fact  
statherin precuro  
beta-amyase (EC 3  
phosphoesterase-re  
hypothetical prote  
probable membrane  
pol protein - simi  
cytochrome-c oxida  
translation initia  
L-ascorbate oxidas  
probable DNA repli  
elongation factor  
P-type ATPase - sl

ALIGNMENTS

RESULT 1  
EHMS  
Ig epsilon chain C region (version 2) - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 16-Aug-1996  
C;Accession: A02145  
R;Ishida, N.; Ueda, S.; Hayashida, H.; Miyata, T.; Honjo, T.  
EMBO J. 1, 1117-1123, 1982  
A;Title: The nucleotide sequence of the mouse immunoglobulin epsilon gene: comparison w  
A;Reference number: A90966, MUID:84236092; PMID:6329728  
A;Accession: A02145  
A;Molecule type: DNA  
A;Residues: 1-423 <ISH>  
A;Note: the sequence was determined from the germline gene  
C;Genetics:  
A;Introns: 91/1; 199/1; 307/1  
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (ka  
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 1  
C;Superfamily: immunoglobulin C region; immunoglobulin homology  
C;Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
F;16-77/Domain: immunoglobulin homology <IMM1>  
F;115-183/Domain: immunoglobulin homology <IMM2>  
F;220-288/Domain: immunoglobulin homology <IMM3>  
F;325-396/Domain: immunoglobulin homology <IMM4>  
F;23-75,122-181,227-286,332-334/Bisulfide bonds: #status predicted  
F;43,84,167,239,262,417/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 82.9%; Score 116; DB 1; Length 423;  
Best Local Similarity 91.3%; Pred. No. 6.9e-10;  
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GYGQSVIVRPPDPKPIVRSITL 24  
Db 281 GYGQCVVDRPDPKPIVRSITL 303

RESULT 2  
I68730  
IgE chain C3 region - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 21-Jan-2000  
C;Accession: I68730  
R;Shinkai, Y.; Nakachi, H.; Honjo, T.; Okumura, K.  
Immunogenetics 27, 288-292, 1988  
A;Title: Mouse immunoglobulin allotypes: Multiple differences between the nucleic acid  
A;Reference number: I54443; MUID:98152907; PMID:3346043  
A;Accession: I68730  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-107 <RES>  
A;Cross-references: GB:M22933; NID:G194464; PIDN:AAA37915.1; PID:G194469  
C;Superfamily: immunoglobulin C region; immunoglobulin homology

APPLICATION NUMBER: US/09/100,414B  
FILING DATE: 20-JUNE-1998  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Maria H. Lin  
REGISTRATION NUMBER: 29,323  
REFERENCE/DOCKET NUMBER: 1151-4157  
TELEPHONE: 212-758-4800  
TELEFAX: 212-751-6849  
INFORMATION FOR SEQ ID NO: 101:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 45 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-100-414B-101

Query Match 71.0%; Score 103; DB 3; Length 45;  
Best Local Similarity 68.0%; Pred. No. 5.1e-09;  
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPPKPIVRSITKC 25  
DB 21 CGETYQSRVTHPHLPALMRSTTKC 45

RESULT 14  
US-09-303-323-101  
Sequence 101, Application US/09303323  
Patent No. 6228987  
GENERAL INFORMATION:  
APPLICANT: Wang, Chang Yi  
TITLE OF INVENTION: NOVEL LHRH PEPTIDE  
TITLE OF INVENTION: IMMUNOGENS  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morgan & Finnegan, L.L.P.  
STREET: 345 Park Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10154-0054  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC Windows  
SOFTWARE: Word 97  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/303,323  
FILING DATE: 30-APR-1999  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/100,414  
FILING DATE: 20-JUNE-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Maria H. Lin  
REGISTRATION NUMBER: 29,323  
REFERENCE/DOCKET NUMBER: 1151-4157  
TELEPHONE: 212-758-4800  
TELEFAX: 212-751-6849  
INFORMATION FOR SEQ ID NO: 101:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 45 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-303-323-101

Query Match 71.0%; Score 103; DB 3; Length 45;  
Best Local Similarity 68.0%; Pred. No. 5.1e-09;  
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPPKPIVRSITKC 25  
DB 21 CGETYQSRVTHPHLPALMRSTTKC 45

RESULT 15  
US-09-770-014-101  
Sequence 101, Application US/09770014  
Patent No. 6559282  
GENERAL INFORMATION:  
APPLICANT: Wang, Chang Yi  
TITLE OF INVENTION: NOVEL LHRH PEPTIDE  
TITLE OF INVENTION: IMMUNOGENS  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morgan & Finnegan, L.L.P.  
STREET: 345 Park Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10154-0054  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC Windows  
SOFTWARE: Word 97  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/770,014  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/100,414  
FILING DATE: 20-JUNE-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Maria H. Lin  
REGISTRATION NUMBER: 29,323  
REFERENCE/DOCKET NUMBER: 1151-4157  
TELEPHONE: 212-758-4800  
TELEFAX: 212-751-6849  
INFORMATION FOR SEQ ID NO: 101:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 45 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-770-014-101

Query Match 71.0%; Score 103; DB 4; Length 45;  
Best Local Similarity 68.0%; Pred. No. 5.1e-09;  
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPPKPIVRSITKC 25  
DB 21 CGETYQSRVTHPHLPALMRSTTKC 45

Search completed: February 26, 2004, 08:29:26  
Job time : 12.6 secs

## TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-758-4800  
TELEFAX: 212-751-6849  
INFORMATION FOR SEQ ID NO: 98:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 42 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-770-014-98

Query Match 71.0%; Score 103; DB 4; Length 42;  
Best Local Similarity 68.0%; Pred. No. 4.8e-09;  
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHHPKPIVRSTKTC 25  
DB 18 CGETYSRVTHPLPALMRSTKTC 42

## RESULT 11

US-09-770-014-99  
Sequence 99, Application US/09770014  
Patent No. 6559282  
GENERAL INFORMATION:  
APPLICANT: Wang, Chang Yi  
TITLE OF INVENTION: NOVEL LHRH PEPTIDE  
TITLE OF INVENTION: IMMUNOGENS  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morgan & Finnegan, L.L.P.  
STREET: 345 Park Avenue  
CITY: New York  
STATE: NY

COUNTRY: USA  
ZIP: 10154-0054  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC Windows  
SOFTWARE: Word 97  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/770,014

FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/100,414  
FILING DATE: 20-JUNE-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Maria H. Lin

REGISTRATION NUMBER: 29,323  
REFERENCE/DOCKET NUMBER: 1151-4157  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-758-4800  
TELEFAX: 212-751-6849  
INFORMATION FOR SEQ ID NO: 99:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 42 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-770-014-99

Query Match 71.0%; Score 103; DB 4; Length 42;  
Best Local Similarity 68.0%; Pred. No. 4.8e-09;  
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHHPKPIVRSTKTC 25  
DB 18 CGETYSRVTHPLPALMRSTKTC 42

## RESULT 12

US-09-770-014-100  
Sequence 100, Application US/09770014  
Patent No. 6559282  
GENERAL INFORMATION:

APPLICANT: Wang, Chang Yi  
TITLE OF INVENTION: NOVEL LHRH PEPTIDE  
TITLE OF INVENTION: IMMUNOGENS  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morgan & Finnegan, L.L.P.  
STREET: 345 Park Avenue  
CITY: New York  
STATE: NY

COUNTRY: USA  
ZIP: 10154-0054

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC Windows  
SOFTWARE: Word 97  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/770,014

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/100,414  
FILING DATE: 20-JUNE-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Maria H. Lin  
REGISTRATION NUMBER: 29,323  
REFERENCE/DOCKET NUMBER: 1151-4157  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-758-4800  
TELEFAX: 212-751-6849

INFORMATION FOR SEQ ID NO: 100:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 42 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-770-014-100

Query Match 71.0%; Score 103; DB 4; Length 42;  
Best Local Similarity 68.0%; Pred. No. 4.8e-09;  
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHHPKPIVRSTKTC 25  
DB 18 CGETYSRVTHPLPALMRSTKTC 42

## RESULT 13

US-09-100-414B-101  
Sequence 101, Application US/09100414B  
Patent No. 6025468  
GENERAL INFORMATION:  
APPLICANT: Wang, Chang Yi  
TITLE OF INVENTION: NOVEL LHRH PEPTIDE  
TITLE OF INVENTION: IMMUNOGENS  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morgan & Finnegan, L.L.P.  
STREET: 345 Park Avenue  
CITY: New York  
STATE: NY

COUNTRY: USA  
ZIP: 10154-0054

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC Windows  
SOFTWARE: Word 97  
CURRENT APPLICATION DATA:

```
Query Match 71.0%; Score 103; DB 3; Length 42;
Best Local Similarity 68.0%; Pred. No. 4.8e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYSRVDPHPKPIVRSITKC 25
Db 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 8
US-09-303-323-99
; Sequence 99, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 100:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-303-323-100

Query Match 71.0%; Score 103; DB 3; Length 42;
Best Local Similarity 68.0%; Pred. No. 4.8e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYSRVDPHPKPIVRSITKC 25
Db 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 9
US-09-303-323-100
; Sequence 100, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
```

```
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/303,323
FILING DATE: 30-APR-1999
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/100,414
FILING DATE: 20-JUNE-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 100:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-303-323-100

Query Match 71.0%; Score 103; DB 3; Length 42;
Best Local Similarity 68.0%; Pred. No. 4.8e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYSRVDPHPKPIVRSITKC 25
Db 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 10
US-09-770-014-98
; Sequence 98, Application US/09770014
; Patent No. 6559282
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/770,014
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
```

```

US-09-100-414B-100
; Sequence 100, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/100,414  
FILING DATE: 20-JUNE-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Maxia H. Lin  
REGISTRATION NUMBER: 29,323  
REFERENCE/DOCKET NUMBER: 1151-4157  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-758-4800  
TELEFAX: 212-751-6849  
INFORMATION FOR SEQ ID NO: 98:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 42 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-303-323-98

```
US-09-303-323-95
; Sequence 95, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303.323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100.414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-303-323-95

Query Match 71.0%; Score 103; DB 3; Length 25;
Best Local Similarity 68.0%; Pred. No. 2.7e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHHPKPIVRSITKC 25
Db 1 CGETYQSRVTHPHLPALMRSTKC 25

RESULT 3
US-09-770-014-95
; Sequence 95, Application US/09770014
; Patent No. 655282
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-100-414B-98

Query Match 71.0%; Score 103; DB 3; Length 42;
Best Local Similarity 68.0%; Pred. No. 4.8e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHHPKPIVRSITKC 25
Db 1 CGETYQSRVTHPHLPALMRSTKC 25

RESULT 4
US-09-100-414B-98
; Sequence 98, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-100-414B-98

Query Match 71.0%; Score 103; DB 3; Length 42;
Best Local Similarity 68.0%; Pred. No. 4.8e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
```



CC AAY91228-Y91231 represent CETP-derived peptides and AAY91232-Y91241 are  
CC immunogens comprising a CETP peptide and a Th epitope which may be used  
CC to prevent or treat arteriosclerosis and cardiovascular disease. AAY91247  
CC and AAY91252-Y91257 are HIV-1 neutralising B-cell epitopes, and AAY91248-  
CC Y91251 and AAY91258-Y91273 are antigenic peptides comprising MVH Th and  
CC HIV-1 B-cell epitope which may be used as a component in an anti-HIV-1  
CC vaccine. AAY91198 and AAY91199 are respectively an immunostimulatory  
CC invasion protein epitope from Yersinia species, and hinge spacer peptide,  
CC both of which may optionally be used in the antigenic peptides of the  
CC invention. (Updated on 12-SEP-2003 to standardise OS field)

XX

SQ Sequence 46 AA;

Query Match 100.0%; Score 140; DB 3; Length 46;  
Best Local Similarity 100.0%; Pred. No. 2.9e-14;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYQSRVTHPHLPALMRSTTKC 25

Db 22 CGETYQSRVTHPHLPALMRSTTKC 46

Search completed: February 26, 2004, 08:23:33  
Job time : 48.6 secs



KW breast cancer; endometriosis; boar taint; meat quality; immunocastration.

XX Unidentified.

XX Key Location/Qualifiers

FT Misc-difference 4

FT /label= Ser, Thr

FT Misc-difference 7

FT /label= Lys, Arg

FT Misc-difference 8

FT /label= Gly, Thr

FT Misc-difference 12

FT /label= His, Thr

FT Misc-difference 13

FT /label= Lys, Arg

FT Misc-difference 16

FT /label= Gly, Thr

XX WO9966952-A1.

PN 29-DEC-1999.

XX 21-JUN-1999; 99WO-US013960.

XX 20-JUN-1998; 98US-00100414.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY;

XX WPI; 2000-160562/14.

XX New peptide immunogen containing luteinizing hormone-releasing hormone

XX antigen site and helper T cell epitope, for e.g. contraception and

XX treatment of cancer.

XX Disclosure; Page 93; 102pp; English.

XX The specification describes peptide immunogens comprising a synthetic

XX helper T cell (Th) epitope and a target antigen, luteinizing hormone-

XX releasing hormone (LHRH). The peptide immunogens cause induction of a

XX specific immune response to LHRH which is involved in regulation of

XX spermatogenesis, ovulation, oestrus, sexual development and secretion of

XX sex hormones. Provision of a promiscuous T helper epitope (which is

XX functional in genetically diverse subjects) provides optimum

XX immunogenicity to the B cell epitopes of the target antigen and thus high

XX antibody titres against the target antigen. The peptide immunogens of the

XX invention are used to vaccinate against mammalian LHRH, for use as

XX (reversible) contraceptive; control of hormone-dependent tumours (cancer

XX of prostate or breast, also endometriosis); to prevent boar taint (and

XX improve meat quality) and for immunocastration. The present sequence

XX appears in the specification

XX Sequence 46 AA;

XX Query Match

XX Best Local Similarity 100.0%; Score 140; DB 3; Length 46;

XX Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX

QY 1 CGETYSQSRVTHPLPALMRSTTKC 25

DB 22 CGETYSQSRVTHPLPALMRSTTKC 46

XX

XX RESULT 15

XX ID AAY91213

XX AC AAY91213 standard; peptide; 46 AA.

XX AC AAY91213;

XX 12-SEP-2003 (revised)

XX 22-MAY-2000 (first entry)

XX

XX

XX

XX

XX

XX

XX

XX

DE

XX

XX

KW

KW

KW

KW

KW

KW

KW

KW

XX

XX

OS

OS

OS

OS

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Modified MVF Th epitope/IgE CH3 domain, SEQ ID NO:93.

Promiscuous T-cell epitope; measles virus F protein; MVF;

hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;

luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;

somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;

foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;

Plasmodium falciparum; circumsporozoite; antimalarial; CERP;

cholesteryl ester transport protein; anti-arteriosclerotic.

Measles virus.

Homo sapiens.

Chimeric.

WO9966957-A2.

29-DEC-1999.

21-JUN-1999; 99WO-US013975.

20-JUN-1998; 98US-00100412.

(UNBI-) UNITED BIOMEDICAL INC.

Wang CY;

WPI; 2000-160564/14.

New artificial T helper cell epitope and derived immunogens with target

antigenic site, for immunization against e.g. malaria, arteriosclerosis

or human immune deficiency virus.

Example 6; Page 98; 129pp; English.

The invention relates to novel promiscuous T helper cell epitopes (Th),

and immunogenic peptides comprising the Th epitopes of the invention

along with B cell epitopes. The Th epitopes and peptide immunogens

containing them, are used to induce a T helper cell response,

specifically against Plasmodium falciparum, cholesteryl ester transport

protein (CERP) or HIV epitopes, but more generally against any pathogen,

immunoreactive self-antigen or tumour antigen. The Th epitopes and

peptide immunogens may be used for prevention and/or treatment of

infections (HIV, foot-and-mouth disease or malaria); for cancer

immunotherapy; for inhibition of the action of luteinising hormone

releasing hormone (LHRH) for contraception, treatment of hormone-

dependent cancer, prevention of boar taint in meat, and immunocastration

; for promoting the growth of animals; or for treating allergies or

arteriosclerosis. Incorporation of a promiscuous Th (functional in

genetically diverse subjects) into an immunogen improves capacity to

induce a strong T helper cell-mediated immune response, resulting in

production of antibodies against a target antigen. Th can replace carrier

proteins and pathogen-derived T helper epitopes. Sequence AAY91121

represents a promiscuous T helper epitope from the measles virus F (MVF)

protein and sequences AAY91122-Y91142, AAY91226 and AAY91245-Y91246

represent synthetic Th epitopes based on the MVF Th epitope. Sequence

AAY91143 represents a promiscuous Th epitope from hepatitis B virus (HBV)

surface antigen, and sequences AAY91144-Y91155 are synthetic epitopes

derived from this HBV epitope. AAY91156-Y91196, AAY91227 and AAY91242-

Y91244 are antigenic peptides comprising an LHRH sequence joined to a

promiscuous Th epitope. AAY91197 is the LHRH target antigenic peptide

used in these LHRH antigenic peptides. AAY91200 is somatostatin, and

AAY91201-Y91207 are antigenic peptides comprising somatostatin and a Th

epitope. Somatostatin immunogens may be used to promote growth in

livestock. AAY91208 is a human CD4 CDR2-like domain antigenic site, and

AAY91209-Y90211 are MVF Th epitopes/CD4 CDR2 antigenic peptides which may

be used to prevent HIV infection of T cells. AAY90212 is a modified

version of a human IGE (immunoglobulin E) CH3 domain, and AAY90213-Y90219

are Th epitope/IgE CH3 antigenic peptides which may be used in the

treatment of allergies. AAY91220 is a peptide derived from foot and mouth

disease virus (FMDV) VP1 capsid protein and AAY91221-Y91222 comprise this

peptide and a Th epitope. AAY91223 is a Plasmodium falciparum

circumsporozoite (CS) target antigen, and AAY91224-Y91225 comprise the CS

antigen and an MVF Th epitope and may be used in a malaria vaccine.

RESULT 12  
 AAY80007  
 ID AAY80007 standard; peptide; 45 AA.  
 AC AAY80007;  
 DT 15-MAY-2000 (first entry)  
 XX IGE-CH3 domain antigen peptide antigenic site 15b SEQ ID NO:14.  
 DE Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;  
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;  
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;  
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.  
 XX Unidentified.  
 OS  
 XX WO9967293-A1.  
 FN  
 XX 29-DEC-1999.  
 PD  
 XX 21-JUN-1999; 99WO-US013959.  
 PF  
 XX 20-JUN-1998; 98US-00100287.  
 PR  
 XX (UNBI-) UNITED BIOMEDICAL INC.  
 PA  
 XX Wang CV, Walfield AM;  
 PI  
 XX WPI; 2000-160578/14.  
 DR  
 XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions  
 PT for immunization against allergy.  
 PT  
 XX Claim 14; Page 75; 155pp; English.  
 PS  
 XX The present invention describes immunoglobulin E (IgE)-CH3 domain  
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic and  
 CC anti-asthmatic properties. (I) induces polyclonal antibodies specific for  
 CC a target effector site on the epsilon-heavy chain of IgE, and so  
 CC preventing triggering and activation of mast cells and basophils and  
 CC downregulation of IgE synthesis. Conjugates, or fusion peptides,  
 CC containing (I) are used for active immunisation against IgE-mediated  
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy  
 CC dermatitis. Nucleic acids that encode these compounds are useful for  
 CC recombinant production of corresponding peptides or in DNA vaccines.  
 CC Conjugates of (I) that include a promiscuous T helper cell epitope  
 CC (functional in genetically diverse subjects), in addition to a B cell  
 CC target epitope, have increased immunogenicity and may include cyclic  
 CC constraints (disulfide bridge) to stabilise conformational features and  
 CC maximize cross-reactivity to the natural target. They induce safe (non-  
 CC anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino acid  
 CC sequences used in the exemplification of the present invention  
 XX  
 SQ Sequence 45 AA;  
 Query Match 100.0%; Score 140; DB 3; Length 45;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-14; Mismatches 0; Gaps 0;  
 Matches 25; Conservative 0; Indels 0;  
 QY 1 CGETYQSRVTHPLPALMRSTTKC 25  
 DB 21 CGETYQSRVTHPLPALMRSTTKC 45  
 RESULT 13  
 ADD89951  
 ID ADD89951 standard; protein; 45 AA.  
 XX  
 AC ADD89951;  
 XX

DT 29-JAN-2004 (first entry)  
 XX IGE peptide used in immunostimulant complex for allergy vaccine.  
 DE  
 XX Immunostimulant; vaccine; human; immunogen; IgE; immunotherapy; allergy;  
 KW antibody; anti-allergic.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FT Modified-site 20 /note= "Epsilon-lysine"  
 PN WO2003068169-A2.  
 XX  
 XX 21-AUG-2003.  
 PD  
 XX 14-FEB-2003; 2003WO-US004711.  
 PF  
 XX 14-FEB-2002; 2002US-00076674.  
 PR  
 XX 31-JAN-2003; 2003US-00076674.  
 PA (UNBI-) UNITED BIOMEDICAL INC.  
 XX  
 XX Sokoll KK;  
 PI  
 XX WPI; 2003-778990/73.  
 DR  
 XX Stabilized immunostimulating complex, useful for vaccination, e.g.  
 PT against human immune deficiency viruses, comprises cationic peptide  
 PT immunogen and anionic oligonucleotide.  
 PT  
 XX Claim 20; SEQ ID NO 11; 159pp; English.  
 PS  
 XX The present sequence is that of a synthetic immunogenic peptide derived  
 CC from human IgE. This is an example of peptides that can be used in  
 CC claimed immunostimulatory complexes of the invention that are  
 CC specifically adapted to act as adjuvant and as peptide immunogen  
 CC stabiliser. The complexes comprise a CpG oligonucleotide and a  
 CC biologically active peptide immunogen. The complex is particulate and can  
 CC efficiently present peptide immunogens to the cells of the immune system  
 CC to produce an immune response. The complexes may be prepared with various  
 CC ratios of peptides to CpG oligonucleotides to provide different physical  
 CC properties, such as the size of the microparticle. An immunostimulatory  
 CC complex comprising the present IgE derived peptide can be used in an anti  
 CC -IGE immunotherapeutic vaccine for the treatment of allergy.  
 XX  
 SQ Sequence 45 AA;  
 Query Match 100.0%; Score 140; DB 7; Length 45;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-14; Mismatches 0; Gaps 0;  
 Matches 25; Conservative 0; Indels 0;  
 QY 1 CGETYQSRVTHPLPALMRSTTKC 25  
 DB 21 CGETYQSRVTHPLPALMRSTTKC 45  
 RESULT 14  
 AAY68603  
 ID AAY68603 standard; peptide; 46 AA.  
 XX  
 AC AAY68603;  
 XX  
 XX 05-MAY-2000 (first entry)  
 DT  
 XX Peptide sequence of the invention.  
 DE  
 XX Helper T cell epitope; peptide immunogen; LHRH;  
 KW luteinising hormone-releasing hormone; spermatogenesis; ovulation;  
 KW oestrus; sexual development; sex hormone; promiscuous T helper epitope;  
 KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;



AC AAY80014;  
 XX  
 DT 15-MAY-2000 (first entry)  
 XX  
 DE IGE immunogenic peptide conjugate SEQ ID NO:21.  
 XX  
 KW Immunoglobulin E; IGE; epsilon heavy chain; antigenic; antigen;  
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;  
 KW antibody; allergic; allergic disease; immunisation; anti-allergic;  
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.  
 XX  
 OS Unidentified.  
 XX  
 PN WO9967293-A1.  
 XX  
 PD 29-DEC-1999.  
 XX  
 PF 21-JUN-1999; 99WO-US013959.  
 XX  
 PR 20-JUN-1998; 98US-00100287.  
 XX  
 PA (UNBI-) UNITED BIOMEDICAL INC.  
 XX  
 PI Wang CY, Walfield AM;  
 XX  
 DR WPI; 2000-160578/14.  
 XX  
 PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions  
 PT for immunization against allergy.  
 XX  
 PS Claim 14; Page 76; 155pp; English.  
 XX  
 CC The present invention describes immunoglobulin E (IGE)-CH3 domain  
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic and  
 CC anti-asthmatic properties. (I) induces polyclonal antibodies specific for  
 CC a target effector site on the epsilon-heavy chain of IGE, and so  
 CC preventing triggering and activation of mast cells and basophils and  
 CC downregulation of IGE synthesis. Conjugates, or fusion peptides,  
 CC containing (I) are used for active immunisation against IGE-mediated  
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy  
 CC dermatitis. Nucleic acids that encode these compounds are useful for  
 CC recombinant production of corresponding peptides or in DNA vaccines.  
 CC Conjugates of (I) that include a promiscuous T helper cell epitope  
 CC (functional in genetically diverse subjects), in addition to a B cell  
 CC target epitope, have increased immunogenicity and may include cyclic  
 CC constraints (disulfide bridge) to stabilise conformational features and  
 CC maximize cross-reactivity to the natural target. They induce safe (non-  
 CC anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino acid  
 XX sequences used in the exemplification of the present invention  
 XX  
 SQ Sequence 42 AA;  
 Query Match 100.0%; Score 140; DB 3; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-14;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGETYQSRVTHPLPALMRSTTKC 25  
 DB 18 CGETYQSRVTHPLPALMRSTTKC 42  
 RESULT 9  
 ADD89950  
 ID ADD89950 standard; protein; 44 AA.  
 XX  
 AC ADD89950;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE IGE peptide used in immunostimulant complex for allergy vaccine.  
 XX  
 KW Immunostimulant; vaccine; human; immunogen; IGE; immunotherapy; allergy;  
 KW antibody; antiallergic.

XX Synthetic.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 19 /note= "Epsilon-lysine"  
 XX  
 PN WO2003068169-A2.  
 XX  
 PD 21-AUG-2003.  
 XX  
 PF 14-FEB-2003; 2003WO-US004711.  
 XX  
 PR 14-FEB-2002; 2002US-00076674.  
 PR 31-JAN-2003; 2003US-00076674.  
 XX  
 PA (UNBI-) UNITED BIOMEDICAL INC.  
 XX  
 PI Sokoll KK;  
 XX  
 DR WPI; 2003-778890/73.  
 XX  
 PT Stabilized immunostimulating complex, useful for vaccination, e.g.  
 PT against human immune deficiency viruses, comprises cationic peptide  
 PT immunogen and anionic oligonucleotide.  
 XX  
 PS Claim 20; SEQ ID NO 10; 159pp; English.  
 XX  
 CC The present sequence is that of a synthetic immunogenic peptide derived  
 CC from human IGE. This is an example of peptides that can be used in  
 CC claimed immunostimulatory complexes of the invention that are  
 CC specifically adapted to act as adjuvant and as peptide immunogen  
 CC stabiliser. The complexes comprise a CpG oligonucleotide and a  
 CC biologically active peptide immunogen. The complex is particulate and can  
 CC efficiently present peptide immunogens to the cells of the immune system  
 CC to produce an immune response. The complexes may be prepared with various  
 CC ratios of peptides to CpG oligonucleotides to provide different physical  
 CC properties, such as the size of the microparticle. An immunostimulatory  
 CC complex comprising the present IGE derived peptide can be used in an anti  
 CC -IGE immunotherapeutic vaccine for the treatment of allergy.  
 XX  
 SQ Sequence 44 AA;  
 Query Match 100.0%; Score 140; DB 7; Length 44;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-14;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGETYQSRVTHPLPALMRSTTKC 25  
 DB 20 CGETYQSRVTHPLPALMRSTTKC 44  
 RESULT 10  
 AAY68605  
 ID AAY68605 standard; peptide; 45 AA.  
 XX  
 AC AAY68605;  
 XX  
 DT 05-MAY-2000 (first entry)  
 XX  
 DE Peptide sequence of the invention.  
 XX  
 KW Helper T cell epitope; peptide immunogen; LHRH;  
 KW luteinising hormone-releasing hormone; spermatogenesis; ovulation;  
 KW oestrus; sexual development; sex hormone; promiscuous T helper epitope;  
 KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;  
 KW breast cancer; endometriosis; boar taint; meat quality; immunocastration.  
 XX  
 OS Unidentified.  
 XX  
 PN WO9966952-A1.  
 XX

derived from this HBV epitope. AAY91156-Y91196, AAY91227 and AAY91242-  
Y91244 are antigenic peptides comprising an LHRH sequence joined to a  
promiscuous Th epitope. AAY91197 is the LHRH target antigenic peptide  
used in these LHRH antigenic peptides. AAY91200 is somatostatin, and  
AAY91201-Y91207 are antigenic peptides comprising somatostatin and a Th  
epitope. Somatostatin immunogens may be used to promote growth in  
livestock. AAY91208 is a human CD4 CDR2-like domain antigenic site, and  
AAY91209-Y90211 are MVH Th epitopes/CD4 CDR2 antigenic peptides which may  
be used to prevent HIV infection of T cells. AAY90212 is a modified  
version of a human IGE (immunoglobulin E) CH3 domain, and AAY90213-Y90219  
are Th epitopes/IGE CH3 antigenic peptides which may be used in the  
treatment of allergies. AAY91220 is a peptide derived from foot and mouth  
disease virus (FMDV) VP1 capsid protein and AAY91221-Y91222 comprise this  
peptide and a Th epitope. AAY91223 is a Plasmodium falciparum  
circumsporozoite (CS) target antigen, and AAY91224-Y91225 comprise the CS  
antigen and an MVF Th epitope and may be used in a malaria vaccine.  
AAY91228-Y91231 represent CERP-derived peptides and AAY91232-Y91241 are  
immunogens comprising a CERP peptide and a Th epitope which may be used  
to prevent or treat arteriosclerosis and cardiovascular disease. AAY91247  
and AAY91252-Y91257 are HIV-1 neutralising B-cell epitopes, and AAY91248-  
Y91251 and AAY91258-Y91273 are antigenic peptides comprising MVH Th and  
HIV-1 B-cell epitope which may be used as a component in an anti-HIV-1  
vaccine. AAY91198 and AAY91199 are respectively an immunostimulatory  
invasin protein epitope from Yersinia species, and hinge spacer peptide,  
both of which may optionally be used in the antigenic peptides of the  
invention. (Updated on 12-SEP-2003 to standardise OS field)

XX Sequence 42 AA;

Query Match 100.0%; Score 140; DB 3; Length 42;  
Best Local Similarity 100.0%; Pred. No. 2.6e-14;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYQSRVTHPHLPALMRSTTKC 25  
DB 18 CGETYQSRVTHPHLPALMRSTTKC 42

RESULT 7

AAY91217  
ID AAY91217 standard; peptide; 42 AA.

XX Homo sapiens.

AC Chimeric.

XX WO9966957-A2.

XX 29-DEC-1999.

DT 12-SEP-2003 (revised)

DT 22-MAY-2000 (first entry)

XX Modified MVF Th epitope/IGE CH3 domain, SEQ ID NO:97.

Promiscuous T-cell epitope; measles virus F protein; MVF;  
hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;  
luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;  
somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;  
foot and mouth disease virus; immunoglobulin E; IGE; anti-allergic;  
Plasmodium falciparum; circumsporozoite; antimalarial; CERP;  
cholesteryl ester transport protein; anti-arteriosclerotic.

XX Measles virus.

OS Homo sapiens.

OS Chimeric.

XX WO9966957-A2.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US013975.

XX 20-JUN-1998; 98US-00100412.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY;

XX

WPI; 2000-160564/14.

DR New artificial T helper cell epitope and derived immunogens with target  
XX antigenic site, for immunization against e.g. malaria, arteriosclerosis  
PT or human immune deficiency virus.  
XX Example 6; Page 99; 129pp; English.

XX The invention relates to novel promiscuous T helper cell epitopes (Th),  
CC and immunogenic peptides comprising the Th epitopes of the invention  
CC along with B cell epitopes. The Th epitopes and peptide immunogens  
CC containing them, are used to induce a T helper cell response,  
CC specifically against Plasmodium falciparum, cholesteryl ester transport  
CC protein (CERP) or HIV epitopes, but more generally against any pathogen,  
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and  
CC peptide immunogens may be used for prevention and/or treatment of  
CC infections (HIV, foot-and-mouth disease or malaria); for cancer  
CC immunotherapy; for inhibition of the action of luteinising hormone  
CC releasing hormone (LHRH) for contraception, treatment of hormone-  
CC dependent cancer, prevention of boar taint in meat, and immunocastration  
CC ; for promoting the growth of animals; or for treating allergies or  
CC arteriosclerosis. Incorporation of a promiscuous Th (functional in  
CC genetically diverse subjects) into an immunogen improves capacity to  
CC induce a strong T helper cell-mediated immune response, resulting in  
CC production of antibodies against a target antigen. Th can replace carrier  
CC proteins and pathogen-derived T helper epitopes. Sequence AAY91121  
CC represents a promiscuous T helper epitope from the measles virus F (MVF)  
CC protein and sequences AAY91122-Y91142, AAY91226 and AAY91245-Y91246  
CC represent synthetic Th epitopes based on the MVF Th epitope. Sequence  
CC AAY91143 represents a promiscuous Th epitope from hepatitis B virus (HBV)  
CC surface antigen, and sequences AAY91144-Y91155 are synthetic epitopes  
CC derived from this HBV epitope. AAY91156-Y91196, AAY91227 and AAY91242-  
CC Y91244 are antigenic peptides comprising an LHRH sequence joined to a  
CC promiscuous Th epitope. AAY91197 is the LHRH target antigenic peptide  
CC used in these LHRH antigenic peptides. AAY91200 is somatostatin, and a Th  
CC AAY91201-Y91207 are antigenic peptides comprising somatostatin, and a Th  
CC epitope. Somatostatin immunogens may be used to promote growth in  
CC livestock. AAY91208 is a human CD4 CDR2-like domain antigenic site, and  
CC AAY91209-Y90211 are MVH Th epitopes/CD4 CDR2 antigenic peptides which may  
CC be used to prevent HIV infection of T cells. AAY90212 is a modified  
CC version of a human IGE (immunoglobulin E) CH3 domain, and AAY90213-Y90219  
CC are Th epitopes/IGE CH3 antigenic peptides which may be used in the  
CC treatment of allergies. AAY91220 is a peptide derived from foot and mouth  
CC disease virus (FMDV) VP1 capsid protein and AAY91221-Y91222 comprise this  
CC peptide and a Th epitope. AAY91223 is a Plasmodium falciparum  
CC circumsporozoite (CS) target antigen, and AAY91224-Y91225 comprise the CS  
CC antigen and an MVF Th epitope and may be used in a malaria vaccine.  
CC immunogens comprising a CERP peptide and a Th epitope which may be used  
CC to prevent or treat arteriosclerosis and cardiovascular disease. AAY91247  
CC and AAY91252-Y91257 are HIV-1 neutralising B-cell epitopes, and AAY91248-  
CC Y91251 and AAY91258-Y91273 are antigenic peptides comprising MVH Th and  
CC HIV-1 B-cell epitope which may be used as a component in an anti-HIV-1  
CC vaccine. AAY91198 and AAY91199 are respectively an immunostimulatory  
CC invasin protein epitope from Yersinia species, and hinge spacer peptide,  
CC both of which may optionally be used in the antigenic peptides of the  
CC invention. (Updated on 12-SEP-2003 to standardise OS field)

XX Sequence 42 AA;

Query Match 100.0%; Score 140; DB 3; Length 42;  
Best Local Similarity 100.0%; Pred. No. 2.6e-14;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYQSRVTHPHLPALMRSTTKC 25  
DB 18 CGETYQSRVTHPHLPALMRSTTKC 42

RESULT 8

AAY90014

ID AAY80014 standard; peptide; 42 AA.

XX

XX	29-DEC-1999.	Matches	25;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
PD											
XX	21-JUN-1999; 99WO-US013975.	QY	1	CGETQSRVTHPHLPALMRSTTKC	25						
XX	20-JUN-1998; 98US-00100412.	DB	18	CGETQSRVTHPHLPALMRSTTKC	42						
XX	(UNBI-) UNITED BIOMEDICAL INC.										
XX	Wang CY;	RESULT 6									
XX	WPI; 2000-160564/14.	AA91215									
XX	New artificial T helper cell epitope and derived immunogens with target	ID	AA91215	standard; peptide; 42 AA.							
XX	or human immune deficiency virus.	XX	AA91215;								
XX	Example 6; Page 99; 129pp; English.	XX	AA91215;								
XX	The invention relates to novel promiscuous T helper cell epitopes (Th),	DT	12-SEP-2003 (revised)								
XX	and immunogenic peptides comprising the Th epitopes of the invention	DT	22-MAY-2000 (first entry)								
XX	along with B cell epitopes. The Th epitopes and peptide immunogens	XX	Modified MWF Th epitope/IgE CH3 domain, SEQ ID NO:95.								
XX	containing them, are used to induce a T helper cell response,	XX	Promiscuous T-cell epitope; measles virus F protein; MWF;								
XX	specifically against Plasmodium falciparum, cholesterol ester transport	XX	hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;								
XX	protein (CTEP) or HIV epitopes, but more generally against any pathogen,	XX	luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;								
XX	immunoreactive self-antigen or tumour antigen. The Th epitopes and	XX	somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;								
XX	peptide immunogens may be used for prevention and/or treatment of	XX	foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;								
XX	disease virus (FMDV) VPI capsid protein and AAY91221-Y91222 comprise this	XX	Plasmodium falciparum; circumsporozoite; antimalarial; CTP;								
XX	peptide and a Th epitope. AAY91223 is a Plasmodium falciparum	XX	cholesterol ester transport protein; anti-arteriosclerotic.								
XX	antigen and an MWF Th epitope and may be used in a malaria vaccine.	XX	Measles virus.								
XX	immunogens comprising a CTEP peptide and a Th epitope which may be used	XX	Homo sapiens.								
XX	to prevent or treat arteriosclerosis and cardiovascular disease. AAY91247	XX	Chimeric.								
XX	and AAY91252-Y91257 are HIV-1 neutralising B-cell epitopes, and AAY91248-	XX	WO9966957-A2.								
XX	Y91251 and AAY91258-Y91273 are antigenic peptides comprising MWF Th and	XX	29-DEC-1999.								
XX	HIV-1 B-cell epitope which may be used as a component in an anti-HIV-1	XX	21-JUN-1999; 99WO-US013975.								
XX	vaccine. AAY91198 and AAY91199 are respectively an immunostimulatory	XX	20-JUN-1998; 98US-00100412.								
XX	invasin protein epitope from <i>Versinia</i> species, and hinge spacer peptide,	XX	(UNBI-) UNITED BIOMEDICAL INC.								
XX	both of which may optionally be used in the antigenic peptides of the	XX	Wang CY;								
XX	invention. (Updated on 12-SEP-2003 to standardise OS field)	XX	WPI; 2000-160564/14.								
XX	Sequence 42 AA;	XX	New artificial T helper cell epitope and derived immunogens with target								
XX	Query Match	XX	antigenic site, for immunization against e.g. malaria, arteriosclerosis								
XX	Best Local Similarity	XX	or human immune deficiency virus.								
XX	100.0%; Score 140; DB 3; Length 42;	XX	Example 6; Page 99; 129pp; English.								
XX	100.0%; Pred. No. 2.6e-14;	XX	The invention relates to novel promiscuous T helper cell epitopes (Th),								
XX		XX	and immunogenic peptides comprising the Th epitopes of the invention								
XX		XX	along with B cell epitopes. The Th epitopes and peptide immunogens								
XX		XX	containing them, are used to induce a T helper cell response,								
XX		XX	specifically against Plasmodium falciparum, cholesterol ester transport								
XX		XX	protein (CTEP) or HIV epitopes, but more generally against any pathogen,								
XX		XX	immunoreactive self-antigen or tumour antigen. The Th epitopes and								
XX		XX	peptide immunogens may be used for prevention and/or treatment of								
XX		XX	infections (HIV, foot-and-mouth disease or malaria); for cancer								
XX		XX	immunotherapy; for inhibition of the action of luteinising hormone								
XX		XX	releasing hormone (LHRH) for contraception, treatment of hormone-								
XX		XX	dependent cancer, prevention of boar taint in meat, and hormone-								
XX		XX	; for promoting the growth of animals; or for treating allergies or								
XX		XX	arteriosclerosis. Incorporation of a promiscuous Th (functional in								
XX		XX	genetically diverse subjects) into an immunogen improves capacity to								
XX		XX	induce a strong T helper cell-mediated immune response, resulting in								
XX		XX	production of antibodies against a target antigen. Th can replace carrier								
XX		XX	proteins and pathogen-derived T helper epitopes. Sequence AAY91121								
XX		XX	represents a promiscuous T helper epitope from the measles virus F (MVF)								
XX		XX	protein and sequences AAY91122-Y91142, AAY91226 and AAY91245-Y91246								
XX		XX	represent synthetic Th epitopes based on the MWF Th epitope. Sequence								
XX		XX	AA91143 represents a promiscuous Th epitope from hepatitis B virus (HBV)								
XX		XX	surface antigen, and sequences AAY91144-Y91155 are synthetic epitopes								
XX		XX	derived from this HBV epitope. AAY91156-Y91196, AAY91227 and AAY91242-								
XX		XX	Y91244 are antigenic peptides comprising an LHRH sequence joined to a								
XX		XX	promiscuous Th epitope. AAY91197 is the LHRH target antigenic peptide								
XX		XX	used in these LHRH antigenic peptides. AAY91200 is somatostatin, and								
XX		XX	AA91201-Y91207 are antigenic peptides comprising somatostatin and a Th								
XX		XX	epitope. Somatostatin immunogens may be used to promote growth in								
XX		XX	livestock. AAY91208 is a human CD4 CDR2-like domain antigenic site, and								
XX		XX	AA91209-Y90211 are MWF Th epitope/CD4 CDR2 antigenic peptides which may								
XX		XX	be used to prevent HIV infection of T cells. AAY90212 is a modified								
XX		XX	version of a human IgE (immunoglobulin E) CH3 domain, and AAY90213-Y90219								
XX		XX	are Th epitope/IgE CH3 antigenic peptides which may be used in the								
XX		XX	treatment of allergies. AAY91220 is a peptide derived from foot and mouth								
XX		XX	disease virus (FMDV) VPI capsid protein and AAY91221-Y91222 comprise this								
XX		XX	peptide and a Th epitope. AAY91223 is a Plasmodium falciparum								
XX		XX	circumsporozoite (CS) target antigen, and AAY91224-Y91225 comprise the CS								
XX		XX	antigen and an MWF Th epitope and may be used in a malaria vaccine.								
XX		XX	immunogens comprising a CTEP peptide and a Th epitope which may be used								
XX		XX	to prevent or treat arteriosclerosis and cardiovascular disease. AAY91247								
XX		XX	and AAY91252-Y91257 are HIV-1 neutralising B-cell epitopes, and AAY91248-								
XX		XX	Y91251 and AAY91258-Y91273 are antigenic peptides comprising MWF Th and								
XX		XX	HIV-1 B-cell epitope which may be used as a component in an anti-HIV-1								
XX		XX	vaccine. AAY91198 and AAY91199 are respectively an immunostimulatory								
XX		XX	invasin protein epitope from <i>Versinia</i> species, and hinge spacer peptide,								
XX		XX	both of which may optionally be used in the antigenic peptides of the								
XX		XX	invention. (Updated on 12-SEP-2003 to standardise OS field)								
XX		XX	Sequence 42 AA;								
XX		XX	Query Match								
XX		XX	Best Local Similarity								
XX		XX	100.0%; Score 140; DB 3; Length 42;								
XX		XX	100.0%; Pred. No. 2.6e-14;								

PI Wang CY, Walfield AM;  
 XX WPI; 2000-160578/14.  
 XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions  
 PT for immunization against allergy.  
 XX  
 PS Claim 1; Page 21; 155pp; English.  
 XX  
 CC The present invention describes immunoglobulin E (IgE)-CH3 domain  
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic and  
 CC anti-asthmatic properties. (I) induces polyclonal antibodies specific for  
 CC a target effector site on the epsilon-heavy chain of IgE, and so  
 CC preventing triggering and activation of mast cells and basophils and  
 CC downregulation of IgE synthesis. Conjugates, or fusion peptides  
 CC containing (I) are used for active immunisation against IgE-mediated  
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy  
 CC dermatitis. Nucleic acids that encode these compounds are useful for  
 CC recombinant production of corresponding peptides or in DNA vaccines.  
 CC Conjugates of (I) that include a promiscuous T helper cell epitope  
 CC (functional in genetically diverse subjects), in addition to a B cell  
 CC target epitope, have increased immunogenicity and may include cyclic  
 CC constraints (disulfide bridge) to stabilise conformational features and  
 CC maximize cross-reactivity to the natural target. They induce safe (non-  
 CC anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino acid  
 CC sequences used in the exemplification of the present invention  
 XX  
 SQ Sequence 25 AA;  
 Query Match 100.0%; Score 140; DB 3; Length 25;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-14;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGETYSQSRVTHPHLPALMRSTTKC 25  
 Db 1 CGETYSQSRVTHPHLPALMRSTTKC 25  
 RESULT 4  
 AAY68604  
 ID AAY68604 standard; peptide; 42 AA.  
 XX  
 AC AAY68604;  
 XX  
 DT 05-MAY-2000 (first entry)  
 XX  
 DE Peptide sequence of the invention.  
 XX  
 KW Helper T cell epitope; peptide immunogen; LHRH;  
 KW luteinising hormone-releasing hormone; spermatogenesis; ovulation;  
 KW oestrus; sexual development; sex hormone; promiscuous T helper epitope;  
 KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;  
 KW breast cancer; endometriosis; boar taint; meat quality; immunocastration.  
 XX  
 OS Unidentified.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1 /label= Ile, Met, Leu  
 FT Misc-difference 2 /label= Ser, Thr  
 FT Misc-difference 5 /label= Lys, Arg  
 FT Misc-difference 6 /label= Gly, Thr  
 FT Misc-difference 10 /label= His, Thr  
 FT Misc-difference 11 /label= Lys, Arg  
 FT Misc-difference 12 /label= Ile, Met, Leu  
 FT Misc-difference 14 /label= Gly, Thr

FT Misc-difference 15 /label= Ile, Met, Val  
 FT  
 XX  
 PN WO9966952-A1.  
 XX  
 XX 29-DEC-1999.  
 XX  
 XX 21-JUN-1999; 99WO-US013960.  
 XX  
 PR 20-JUN-1998; 98US-00100414.  
 XX  
 PA (UNBI-) UNITED BIOMEDICAL INC.  
 XX  
 PI Wang CY;  
 XX  
 DR WPI; 2000-160562/14.  
 XX  
 XX New peptide immunogen containing luteinising hormone-releasing hormone  
 PT antigen site and helper T cell epitope, for e.g. contraception and  
 PT treatment of cancer.  
 XX  
 PS Disclosure; Page 95; 102pp; English.  
 XX  
 CC The specification describes peptide immunogens comprising a synthetic  
 CC helper T cell (Th) epitope and a target antigen, luteinising hormone-  
 CC releasing hormone (LHRH). The peptide immunogens cause induction of a  
 CC specific immune response to LHRH which is involved in regulation of  
 CC spermatogenesis, ovulation, oestrus, sexual development and secretion of  
 CC sex hormones. Provision of a promiscuous T helper epitope (which is  
 CC functional in genetically diverse subjects) provides optimum  
 CC immunogenicity to the B cell epitopes of the target antigen and thus high  
 CC antibody titres against the target antigen. The peptide immunogens of the  
 CC invention are used to vaccinate against mammalian LHRH, for use as  
 CC (reversible) contraceptive; control of hormone-dependent tumours (cancer  
 CC of prostate or breast, also endometriosis); to prevent boar taint (and  
 CC improve meat quality) and for immunocastration. The present sequence  
 CC appears in the specification  
 XX  
 SQ Sequence 42 AA;  
 Query Match 100.0%; Score 140; DB 3; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-14;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGETYSQSRVTHPHLPALMRSTTKC 25  
 Db 18 CGETYSQSRVTHPHLPALMRSTTKC 42  
 RESULT 5  
 AAY91216  
 ID AAY91216 standard; peptide; 42 AA.  
 XX  
 AC AAY91216;  
 XX  
 DT 12-SEP-2003 (revised)  
 DT 22-MAY-2000 (first entry)  
 XX  
 DE Modified MVP Th epitope/IgE CH3 domain, SEQ ID NO:96.  
 XX  
 KW Promiscuous T-cell epitope; measles virus F protein; MVF;  
 KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;  
 KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;  
 KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;  
 KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;  
 KW Plasmodium falciparum; circumsporozoite; antimalarial; CTRP;  
 KW cholesteryl ester transport protein; anti-arteriosclerotic.  
 XX  
 OS Measles virus.  
 OS Homo sapiens.  
 OS Chimeric.  
 XX  
 PN WO9966957-A2.

CC (reversible) contraceptive; control of hormone-dependent tumours (cancer  
 CC of prostate or breast; also endometriosis); to prevent boar taint (and  
 CC improve meat quality) and for immunocastration. The present sequence  
 CC appears in the specification  
 XX  
 SQ Sequence 25 AA;  
 Query Match 100.0%; Score 140; DB 3; Length 25;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-14; Indels 0; Gaps 0;  
 Matches 25; Conservative 0; Mismatches 0;  
 QY 1 CGETYQSRVTHPHLPALMRSTTKC 25  
 DB 1 CGETYQSRVTHPHLPALMRSTTKC 25  
 RESULT 2  
 ID AAY91212 standard; peptide; 25 AA.  
 AC AAY91212;  
 DT 22-MAY-2000 (first entry)  
 DE Modified human IGE CH3 domain, SEQ ID NO:92.  
 XX Promiscuous T-cell epitope; measles virus F protein; MVF;  
 KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;  
 KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;  
 KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;  
 KW foot and mouth disease virus; immunoglobulin E; IGE; anti-allergic;  
 KW Plasmodium falciparum; circumsporozoite; antimalarial; CERP;  
 KW cholesterol ester transport protein; anti-arteriosclerotic.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX WO9966957-A2.  
 DN 29-DEC-1999.  
 PD 21-JUN-1999; 99WO-US013975.  
 PF 20-JUN-1998; 98US-00100412.  
 PR (UNBI-) UNITED BIOMEDICAL INC.  
 PA Wang CY;  
 XX WPI; 2000-160564/14.  
 DR New artificial T helper cell epitope and derived immunogens with target  
 PT antigenic site, for immunization against e.g. malaria, arteriosclerosis  
 PT or human immune deficiency virus.  
 XX  
 PS Example 6; Page 40; 129pp; English.  
 XX The invention relates to novel promiscuous T helper cell epitopes (Th),  
 CC and immunogenic peptides comprising the Th epitopes of the invention  
 CC along with B cell epitopes. The Th epitopes and peptide immunogens  
 CC containing them, are used to induce a T helper cell response,  
 CC specifically against Plasmodium falciparum, cholesterol ester transport  
 CC protein (CERP) or HIV epitopes, but more generally against any pathogen,  
 CC immunoreactive self-antigen or tumour antigen. The Th epitopes and  
 CC peptide immunogens may be used for prevention and/or treatment of  
 CC infections (HIV, foot-and-mouth disease or malaria); for cancer  
 CC immunotherapy; for inhibition of the action of luteinising hormone  
 CC releasing hormone (LHRH) for contraception, treatment of hormone-  
 CC dependent cancer, prevention of boar taint in meat, and immunocastration  
 CC ; for promoting the growth of animals; or for treating allergies or  
 CC arteriosclerosis. Incorporation of a promiscuous Th (functional in  
 CC genetically diverse subjects) into an immunogen improves capacity to  
 CC induce a strong T helper cell-mediated immune response, resulting in

CC production of antibodies against a target antigen. Th can replace carrier  
 CC proteins and pathogen-derived T helper epitopes. Sequence AAY91212  
 CC represents a promiscuous T helper epitope from the measles virus F (VNF)  
 CC protein and sequences AAY91212-Y91142, AAY91226 and AAY91245-Y91246  
 CC represent synthetic Th epitopes based on the MVF Th epitope. Sequence  
 CC AAY91143 represents a promiscuous Th epitope from hepatitis B virus (HBV)  
 CC surface antigen, and sequences AAY91144-Y91155 are synthetic epitopes  
 CC derived from this HBV epitope. AAY91156-Y91196, AAY91227 and AAY91242-  
 CC Y91244 are antigenic peptides comprising an LHRH sequence joined to a  
 CC promiscuous Th epitope. AAY91197 is the LHRH target antigenic peptide  
 CC used in these LHRH antigenic peptides. AAY91200 is somatostatin, and  
 CC AAY91201-Y91207 are antigenic peptides comprising somatostatin and a Th  
 CC epitope. Somatostatin immunogens may be used to promote growth in  
 CC livestock. AAY91208 is a human CD4 CDR2-like domain antigenic site, and  
 CC AAY91209-Y90211 are MVF Th epitope/CD4 CDR2 antigenic peptides which may  
 CC be used to prevent HIV infection of T cells. AAY90212 is a modified  
 CC version of a human IGE (immunoglobulin E) CH3 domain, and AAY90213-Y90219  
 CC are Th epitope/IGE CH3 antigenic peptides which may be used in the  
 CC treatment of allergies. AAY91220 is a peptide derived from foot and mouth  
 CC disease virus (FMDV) VP1 capsid protein and AAY91221-Y91222 comprise this  
 CC peptide and a Th epitope. AAY91223 is a Plasmodium falciparum  
 CC circumsporozoite (CS) target antigen, and AAY91224-Y91225 comprise the CS  
 CC antigen and an MVF Th epitope and may be used in a malaria vaccine.  
 CC AAY91228-Y91231 represent CERP-derived peptides and AAY91232-Y91241 are  
 CC immunogens comprising a CERP peptide and a Th epitope which may be used  
 CC to prevent or treat arteriosclerosis and cardiovascular disease. AAY91247  
 CC and AAY91252-Y91257 are HIV-1 neutralising B-cell epitopes, and AAY91248-  
 CC Y91251 and AAY91258-Y91273 are antigenic peptides comprising MVF Th and  
 CC HIV-1 B-cell epitope which may be used as a component in an anti-HIV-1  
 CC vaccine. AAY91198 and AAY91199 are respectively an immunostimulatory  
 CC invasin protein epitope from Yersinia species, and hinge spacer peptide,  
 CC both of which may optionally be used in the antigenic peptides of the  
 CC invention  
 XX  
 SQ Sequence 25 AA;

Query Match 100.0%; Score 140; DB 3; Length 25;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-14; Indels 0; Gaps 0;  
 Matches 25; Conservative 0; Mismatches 0;

QY 1 CGETYQSRVTHPHLPALMRSTTKC 25  
 DB 1 CGETYQSRVTHPHLPALMRSTTKC 25

RESULT 3  
 ID AAY79998 standard; peptide; 25 AA.  
 AC AAY79998;  
 DT 15-MAY-2000 (first entry)  
 DE Optimised IGE-CH3 domain antigen peptide for human IGE.

XX Immunoglobulin E; IGE; epsilon heavy chain; antigenic; antigen;  
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;  
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;  
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX WO9967293-A1.  
 PD 29-DEC-1999.  
 XX 21-JUN-1999; 99WO-US013959.  
 PF 20-JUN-1998; 98US-00100287.  
 PR (UNBI-) UNITED BIOMEDICAL INC.



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 08:16:50 ; Search time 46.6 Seconds  
(without alignments)  
151.581 Million cell updates/sec

Title: US-09-701-623C-5

Perfect score: 140

Sequence: 1 CGETYQSRVTHPLFRALMRSTTKC 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	140	100.0	25	3 AAY68602	Aay68602 Peptide s
2	140	100.0	25	3 AAY91212	Aay91212 Modified
3	140	100.0	25	3 AAY79998	Aay79998 Optimised
4	140	100.0	42	3 AAY68604	Aay68604 Peptide s
5	140	100.0	42	3 AAY91216	Aay91216 Modified
6	140	100.0	42	3 AAY91215	Aay91215 Modified
7	140	100.0	42	3 AAY91217	Aay91217 Modified
8	140	100.0	42	3 AAY80014	Aay80014 Ige immu
9	140	100.0	44	7 ADD89950	Add89950 Ige pepti
10	140	100.0	45	3 AAY68605	Aay68605 Peptide s
11	140	100.0	45	3 AAY91218	Aay91218 Modified
12	140	100.0	45	3 AAY80007	Aay80007 Ige-CH3 d
13	140	100.0	45	7 ADD89951	Add89951 Ige pepti
14	140	100.0	46	3 AAY68603	Aay68603 Peptide s
15	140	100.0	46	3 AAY91213	Aay91213 Modified
16	140	100.0	46	3 AAY91214	Aay91214 Modified
17	140	100.0	46	3 AAY80011	Aay80011 Ige immu
18	140	100.0	56	3 AAY80016	Aay80016 Ige immu
19	140	100.0	59	3 AAY80010	Aay80010 Ige immu
20	140	100.0	60	3 AAY80013	Aay80013 Ige immu
21	140	100.0	60	3 AAY80015	Aay80015 Ige immu
22	140	100.0	63	3 AAY68606	Aay68606 Peptide s
23	140	100.0	63	3 AAY91219	Aay91219 Inv epit
24	140	100.0	63	3 AAY80008	Aay80008 Ige-CH3 d
25	140	100.0	63	3 AAY80012	Aay80012 Ige immu

26	133	95.0	60	3 AAY800078	Aay800078 Ige immu
27	117	83.6	106	2 AAY42620	Aay42620 Human IGB
28	117	83.6	110	2 AAR32034	Aar32034 Variant I
29	117	83.6	110	2 AAR33328	Aar33328 Variant I
30	117	83.6	110	2 AAR33321	Aar33321 Variant I
31	117	83.6	110	2 AAR33322	Aar33322 Variant I
32	117	83.6	110	2 AAR33681	Aar33681 Variant I
33	117	83.6	110	2 AAR33682	Aar33682 Variant I
34	117	83.6	110	2 AAR32025	Aar32025 Variant I
35	117	83.6	110	2 AAR32026	Aar32026 Variant I
36	117	83.6	110	2 AAR33330	Aar33330 Variant I
37	117	83.6	110	2 AAR33680	Aar33680 Variant I
38	117	83.6	110	2 AAR33691	Aar33691 Variant I
39	117	83.6	110	2 AAR32033	Aar32033 Variant I
40	117	83.6	110	2 AAR33319	Aar33319 Variant I
41	117	83.6	110	2 AAR33678	Aar33678 Variant I
42	117	83.6	110	2 AAR35744	Aar35744 Variant I
43	117	83.6	110	2 AAR35747	Aar35747 Variant I
44	117	83.6	110	2 AAR35748	Aar35748 Variant I
45	117	83.6	110	2 AAR35751	Aar35751 Variant I

ALIGNMENTS

RESULT 1  
AAY68602  
ID AAY68602 standard; peptide; 25 AA.  
XX  
AC AAY68602;  
XX  
DT 05-MAY-2000 (first entry)  
XX  
DE Peptide sequence of the invention.  
XX  
KW Helper T cell epitope; peptide immunogen; LHRH;  
KW luteinising hormone-releasing hormone; spermatogenesis; ovulation;  
KW oestrus; sexual development; sex hormone; promiscuous T helper epitope;  
KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;  
KW breast cancer; endometriosis; boar taint; meat quality; immunocastration.  
XX  
OS Unidentified.  
XX  
PN WO9966952-A1.  
XX  
PD 29-DEC-1999.  
XX  
PF 21-JUN-1999; 99WO-US013960.  
XX  
PR 20-JUN-1998; 98US-00100414.  
XX  
PA (UNBI-) UNITED BIOMEDICAL INC.  
XX  
PI Wang CY;  
XX  
DR WPI; 2000-160562/14.  
XX  
PT New peptide immunogen containing luteinizing hormone-releasing hormone  
PT antigen site and helper T cell epitope, for e.g. contraception and  
PT treatment of cancer.  
XX  
FS Disclosure; Page 92; 102pp; English.  
XX  
CC The specification describes peptide immunogens comprising a synthetic  
CC helper T cell (Th) epitope and a target antigen, luteinising hormone-  
CC releasing hormone (LHRH). The peptide immunogens cause induction of a  
CC specific immune response to LHRH which is involved in regulation of  
CC spermatogenesis, ovulation, oestrus, sexual development and secretion of  
CC sex hormones. Provision of a promiscuous T helper epitope (which is  
CC functional in genetically diverse subjects) provides optimum  
CC immunogenicity to the B cell epitopes of the target antigen and thus high  
CC antibody titres against the target antigen. The peptide immunogens of the  
CC invention are used to vaccinate against mammalian LHRH, for use as

CC preventing IgE from binding to high affinity receptors on mast cells and  
 CC basophils. The products of the invention are useful in the manufacture of  
 CC a medicament for treating or preventing IgE-mediated allergic disorders  
 CC including asthma, allergic rhinitis, gastrointestinal allergies such as  
 CC food allergies, eosinophilia, conjunctivitis, glomerular nephritis, flea  
 CC allergies or atopic dermatitis, in an animal, e.g. human or dog. The  
 CC polynucleotide products are useful for treating IgE-mediated allergic  
 CC disorders, by gene therapy. Antigenic peptides comprising conserved amino  
 CC acid residues of the CH3 domain of an IgE molecule from one species  
 CC flanked by variable amino acid residues of the CH3 domain of an IgE  
 CC molecule from a second unrelated species are capable of inducing a high  
 CC titre of anti-IgE antibodies when administered to an animal without  
 CC causing anaphylaxis. ABG74772-ABG74785 represents polypeptide sequences  
 CC used in designing the constructs described in the disclosure of the  
 CC invention  
 XX  
 SQ Sequence 346 AA;

Query Match 83.7%; Score 118; DB 6; Length 346;  
 Best Local Similarity 95.7%; Pred. No. 1.7e-09;  
 Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GETYSRVTHPLPKDIVRSIAK 24  
 DB 212 GETYCRVTHPLPKDIVRSIAK 234

## RESULT 15

ABG74783  
 ID ABG74783 standard; protein; 347 AA.

XX  
 AC ABG74783;

XX  
 DT 05-JUN-2003 (first entry)

XX  
 DE Human CH2-human/canine CH3-human CH4 IgE-3 fusion protein.

XX CH3 domain; IgE; antigen; non-anaphylactic; anti-IgE; fusion protein;  
 KW dermatological; antiinflammatory; ophthalmological; allergy; asthma;  
 KW allergic rhinitis; gastrointestinal allergy; food allergy; eosinophilia;  
 KW conjunctivitis; glomerular nephritis; flea allergy; atopic dermatitis;  
 KW gene therapy; human; canine; dog.

OS Homo sapiens.  
 OS Canis familiaris.

XX  
 PN EP1262491-A2.

XX  
 PD 04-DEC-2002.

XX  
 PF 22-MAY-2002; 2002EP-00253606.

XX  
 PR 22-MAY-2001; 2001US-0292638P.

XX  
 PA (PFIZ ) PFIZER PROD INC.

XX  
 PI Brown TM, Morsey MA;

XX  
 XX WPI; 2003-122561/12.

DR  
 DR N-PSDB; ACA55181.

XX  
 PT Novel isolated antigenic peptide comprising amino acid residues of CH3  
 PT domain of IgE molecule from first species and a second unrelated species,  
 PT induces non-anaphylactic anti-IgE immune response in animal.

XX  
 PS Claim 3; Page 37-39; 50pp; English.

XX This invention describes a novel antigenic peptide comprising amino acid  
 CC residues of an IgE CH3 domain from a first species (ADE1) and amino acid  
 CC residues of an IgE CH3 domain of a second unrelated species (ADE2), where  
 CC ADE1 is conserved in the IgE CH3 domain of the second species and ADE2 is  
 CC not conserved in the IgE CH3 domain of the first species. The novel  
 CC antigenic peptide induces a non-anaphylactic anti-IgE immune response in

CC an animal. The invention also discloses the polynucleotide sequence  
 CC encoding the antigenic peptide and an antigenic fusion protein comprising  
 CC the antigenic peptide of the invention and a heterologous protein  
 CC carrier, where the fusion protein induces an anti-IgE immune response  
 CC that does not cause anaphylaxis when administered to an animal. The  
 CC products of the invention have dermatological, antiinflammatory and  
 CC ophthalmological activity. The antigenic peptide described is capable of  
 CC preventing IgE from binding to high affinity receptors on mast cells and  
 CC basophils. The products of the invention are useful in the manufacture of  
 CC a medicament for treating or preventing IgE-mediated allergic disorders  
 CC including asthma, allergic rhinitis, gastrointestinal allergies such as  
 CC food allergies, eosinophilia, conjunctivitis, glomerular nephritis, flea  
 CC allergies or atopic dermatitis, in an animal, e.g. human or dog. The  
 CC polynucleotide products are useful for treating IgE-mediated allergic  
 CC disorders, by gene therapy. Antigenic peptides comprising conserved amino  
 CC acid residues of the CH3 domain of an IgE molecule from one species  
 CC flanked by variable amino acid residues of the CH3 domain of an IgE  
 CC molecule from a second unrelated species are capable of inducing a high  
 CC titre of anti-IgE antibodies when administered to an animal without  
 CC causing anaphylaxis. ABG74772-ABG74785 represents polypeptide sequences  
 CC used in designing the constructs described in the disclosure of the  
 CC invention

XX  
 SQ Sequence 347 AA;

Query Match 83.7%; Score 118; DB 6; Length 347;  
 Best Local Similarity 95.7%; Pred. No. 1.7e-09;  
 Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GETYSRVTHPLPKDIVRSIAK 24

DB 213 GETYCRVTHPLPKDIVRSIAK 235

Search completed: February 26, 2004, 08:23:34  
 Job time : 47.6 secs

PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions  
 PT for immunization against allergy.  
 XX  
 PS Example 1; Page 66-68; 155pp; English.  
 XX  
 CC The present invention describes immunoglobulin E (IgE)-CH3 domain  
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic and  
 CC anti-asthmatic properties. (I) induces polyclonal antibodies specific for  
 CC a target effector site on the epsilon-heavy chain of IgE, and so  
 CC preventing triggering and activation of mast cells and basophils and  
 CC downregulation of IgE synthesis. Conjugates, or fusion peptides,  
 CC containing (I) are used for active immunisation against IgE-mediated  
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy  
 CC dermatitis. Nucleic acids that encode these compounds are useful for  
 CC recombinant production of corresponding peptides or in DNA vaccines.  
 CC Conjugates of (I) that include a promiscuous T helper cell epitope  
 CC (functional in genetically diverse subjects), in addition to a B cell  
 CC target epitope, have increased immunogenicity and may include cyclic  
 CC constraints (disulfide bridge) to stabilise conformational features and  
 CC maximize cross-reactivity to the natural target. They induce safe (non-  
 CC anaphylactogenic) antibodies. AAY7994 to AAY8004 represent amino acid  
 CC sequences used in the exemplification of the present invention  
 XX  
 SQ Sequence 312 AA;  
 Query Match 83.7%; Score 118; DB 3; Length 312;  
 Best Local Similarity 95.7%; Pred. No. 1.5e-09;  
 Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 GETYYSRVTHPLPKDIVRSIAK 24  
 DB 187 GETYYCRVTHPLPKDIVRSIAK 209  
 RESULT 13  
 AAB06208  
 ID AAB06208 standard; protein; 341 AA.  
 XX  
 AC AAB06208;  
 XX  
 DT 12-SEP-2003 (revised)  
 DT 22-NOV-2000 (first entry)  
 XX  
 DE Immunogenic peptide consisting of opossum CH2, dog CH3 and opossum CH4.  
 XX  
 KW Dog; opossum; immunoglobulin E; IgE; vaccination; infection; allergy;  
 KW asthma; eczema; immunogenic peptide.  
 XX  
 OS Didelphis virginiana.  
 OS Canis sp.  
 OS Chimeric.  
 XX  
 PN WO200025722-A2.  
 XX  
 PD 11-MAY-2000.  
 XX  
 PF 21-OCT-1999; 99WO-SE001896.  
 XX  
 PR 02-NOV-1998; 98US-0106652P.  
 PR 22-SEP-1999; 99US-00401636.  
 XX  
 XX (RESI-) RESISTENTIA PHARM AB.  
 XX  
 PI Hellman LT;  
 XX  
 DR WPI; 2000-365342/31.  
 XX  
 XX Immunogenic polypeptides useful for preventing the harmful effects of  
 PT immunoglobulin E in mammals.  
 PT  
 XX Disclosure; Fig 2; 50pp; English.  
 PS  
 XX The present sequence is an immunogenic peptide consisting of the heavy

CC chain constant regions 2 and 4 of the opossum IgE and the heavy chain  
 CC constant region 3 from the dog. It was shown to cause a stronger  
 CC polyclonal anti-self IgE response than peptides consisting of the same  
 CC regions from one mammal. Immunogenic peptides, particularly those  
 CC consisting of different heavy chain constant regions, can be used for  
 CC vaccination in humans, against bacterial and viral infections and  
 CC allergies, such as asthma, fur, pollen and food allergies and eczema.  
 CC (Updated on 12-SEP-2003 to standardise OS field)  
 XX  
 SQ Sequence 341 AA;  
 Query Match 83.7%; Score 118; DB 3; Length 341;  
 Best Local Similarity 95.7%; Pred. No. 1.7e-09;  
 Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 GETYYSRVTHPLPKDIVRSIAK 24  
 DB 206 GETYYCRVTHPLPKDIVRSIAK 228  
 RESULT 14  
 ABG74781  
 ID ABG74781 standard; protein; 346 AA.  
 XX  
 AC ABG74781;  
 XX  
 DT 05-JUN-2003 (first entry)  
 DT  
 XX  
 DE Human IgE CH2-canine CH3-human CH4 IgE-1 fusion protein.  
 XX  
 KW CH3 domain; IgE; antigen; non-anaphylactic; anti-IgE; fusion protein;  
 KW dermatological; antiinflammatory; ophthalmological; allergy; asthma;  
 KW allergic rhinitis; gastrointestinal allergy; food allergy; eosinophilia;  
 KW conjunctivitis; glomerular nephritis; flea allergy; atopic dermatitis;  
 KW gene therapy; human; canine; dog.  
 XX  
 OS Homo sapiens.  
 OS Canis familiaris.  
 XX  
 PN EP1262491-A2.  
 XX  
 PD 04-DEC-2002.  
 XX  
 PF 22-MAY-2002; 2002EP-00253606.  
 XX  
 PR 22-MAY-2001; 2001US-0292638P.  
 XX  
 PA (PFI2 ) PFIZER PROD INC.  
 XX  
 XX Brown TM, Morse MA;  
 XX  
 XX WPI; 2003-122561/12.  
 XX N-PSDB; ACA55179.  
 DR  
 DR  
 XX Novel isolated antigenic peptide comprising amino acid residues of CH3  
 PT domain of IgE molecule from first species and a second unrelated species,  
 PT induces non-anaphylactic anti-IgE immune response in animal.  
 XX  
 XX Claim 3; Page 34-36; 50pp; English.  
 PS  
 XX This invention describes a novel antigenic peptide comprising amino acid  
 CC residues of an IgE CH3 domain from a first species (ADE1) and amino acid  
 CC residues of an IgE CH3 domain of a second unrelated species (ADE2), where  
 CC ADE1 is conserved in the IgE CH3 domain of the second species and ADE2 is  
 CC not conserved in the IgE CH3 domain of the first species. The novel  
 CC antigenic peptide induces a non-anaphylactic anti-IgE immune response in  
 CC an animal. The invention also discloses the polynucleotide sequence  
 CC encoding the antigenic peptide and an antigenic fusion protein comprising  
 CC the antigenic peptide of the invention and a heterologous protein  
 CC carrier, where the fusion protein induces an anti-IgE immune response  
 CC that does not cause anaphylaxis when administered to an animal. The  
 CC products of the invention have dermatological, antiinflammatory and  
 CC ophthalmological activity. The antigenic peptide described is capable of

```

XX OS Homo sapiens.
XX OS Canis familiaris.
XX OS Synthetic.
XX PN EP1262491-A2.
XX PD 04-DEC-2002.
XX XX
XX PF 22-MAY-2002; 2002EP-00253606.
XX XX
XX PR 22-MAY-2001; 2001US-0292638P.
XX XX
XX PA (PFIZ ) PFIZER PROD INC.
XX XX
XX PI Brown TM, Morsey MA;
XX XX
XX DR WPI; 2003-122561/12.
XX DR N-PSDB; ACA55171.
XX XX
XX PT Novel isolated antigenic peptide comprising amino acid residues of CH3
XX PT domain of IgE molecule from first species and a second unrelated species,
XX PT induces non-anaphylactic anti-IgE immune response in animal.
XX XX
XX PS Claim 2; Page 29; 50pp; English.
XX XX
XX CC This invention describes a novel antigenic peptide comprising amino acid
XX CC residues of an IgE CH3 domain from a first species (ADE1) and amino acid
XX CC residues of an IgE CH3 domain of a second unrelated species (ADE2), where
XX CC ADE1 is conserved in the IgE CH3 domain of the second species and ADE2 is
XX CC not conserved in the IgE CH3 domain of the first species. The novel
XX CC antigenic peptide induces a non-anaphylactic anti-IgE immune response in
XX CC an animal. The invention also discloses the polynucleotide sequence
XX CC encoding the antigenic peptide and an antigenic fusion protein comprising
XX CC the antigenic peptide of the invention and a heterologous protein
XX CC carrier, where the fusion protein induces an anti-IgE immune response
XX CC that does not cause anaphylaxis when administered to an animal. The
XX CC products of the invention have dermatological, antiinflammatory and
XX CC ophthalmological activity. The antigenic peptide described is capable of
XX CC preventing IgE from binding to high affinity receptors on mast cells and
XX CC basophils. The products of the invention are useful in the manufacture of
XX CC a medicament for treating or preventing IgE-mediated allergic disorders
XX CC including asthma, allergic rhinitis, gastrointestinal allergies such as
XX CC food allergies, eosinophilia, conjunctivitis, glomerular nephritis, flea
XX CC allergies or atopic dermatitis, in an animal, e.g. human or dog. The
XX CC polynucleotide products are useful for treating IgE-mediated allergic
XX CC disorders, by gene therapy. Antigenic peptides comprising conserved amino
XX CC acid residues of the CH3 domain of an IgE molecule from one species
XX CC flanked by variable amino acid residues of the CH3 domain of an IgE
XX CC molecule from a second unrelated species are capable of inducing a high
XX CC titre of anti-IgE antibodies when administered to an animal without
XX CC causing anaphylaxis. ABG74772-ABG74785 represents polypeptide sequences
XX CC used in designing the constructs described in the disclosure of the
XX CC invention
XX SQ Sequence 117 AA;

Query Match 83.7%; Score 118; DB 6; Length 117;
Best Local Similarity 95.7%; Pred. No. 4.9e-10;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GETTYSRVTTHPLPKDIVRSIAK 24
DB 87 GETTYCRVTHPLPKDIVRSIAK 109

RESULT 11
AAW24097
ID AAW24097 standard; peptide; 124 AA.
XX AC
XX AC AAW24097;
XX XX
XX DT 21-NOV-1997 (first entry)

Partial canine immunoglobulin E protein.
Immunoglobulin E; IgE; anti-canine IgE antibody; allergy; canine; dog.
Canis familiaris.
JP09169795-A.
30-JUN-1997.
22-DEC-1995; 95JP-00334381.
22-DEC-1995; 95JP-00334381.
(HITB ) HITACHI CHEM CO LTD.
WPI; 1997-389423/36.
N-PSDB; AAI85646.
Canine immunoglobulin E peptide fragment and related DNA - useful for the
Preparation of anti-canine immunoglobulin E antibody.
Claim 1; Page 8; 12pp; Japanese.
This is a partial canine immunoglobulin E (IgE) protein. Peptide
fragments (AAW24098-106) containing at least five continuous amino acids
of this sequence are used for the preparation of anti-canine IgE
antibody. The anti-canine IgE antibody can be used for the diagnosis of
canine allergies
SQ Sequence 124 AA;

Query Match 83.7%; Score 118; DB 2; Length 124;
Best Local Similarity 95.7%; Pred. No. 5.2e-10;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GETTYSRVTTHPLPKDIVRSIAK 24
DB 57 GETTYCRVTHPLPKDIVRSIAK 79

RESULT 12
AAW79995
ID AAW79995 standard; protein; 312 AA.
XX AC
XX AC AAY79995;
XX DT
XX DT 15-MAY-2000 (first entry)
XX DE
XX DE Dog immunoglobulin E epsilon heavy chain SEQ ID NO:2.
XX KW
XX KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
XX KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
XX KW antibody; allergy; allergic disease; immunisation; anti-allergic;
XX KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
XX OS
XX OS Canis sp.
XX XX
XX XX WO9967293-A1.
XX XX
XX PD 29-DEC-1999.
XX XX
XX PF 21-JUN-1999; 99WO-US013959.
XX XX
XX PR 20-JUN-1998; 98US-00100287.
XX XX
XX XX (UNBI-) UNITED BIOMEDICAL INC.
XX XX
XX PI Wang CY, Walfield AM;
XX XX
XX DR WPI; 2000-160578/14.

```

DR WPI; 2003-122561/12.  
 DR N-PSDB; ACA55170.  
 XX Novel isolated antigenic peptide comprising amino acid residues of CH3  
 PT domain of IGE molecule from first species and a second unrelated species,  
 PT induces non-anaphylactic anti-IgE immune response in animal.  
 XX  
 PS Claim 15; Page 28; 50pp; English.  
 XX  
 CC This invention describes a novel antigenic peptide comprising amino acid  
 CC residues of an IGE CH3 domain from a first species (ADE1) and amino acid  
 CC residues of an IGE CH3 domain of a second unrelated species (ADE2), where  
 CC ADE1 is conserved in the IGE CH3 domain of the second species and ADE2 is  
 CC not conserved in the IGE CH3 domain of the first species. The novel  
 CC antigenic peptide induces a non-anaphylactic anti-IgE immune response in  
 CC an animal. The invention also discloses the polynucleotide sequence  
 CC encoding the antigenic peptide and an antigenic fusion protein comprising  
 CC the antigenic peptide of the invention and a heterologous protein  
 CC carrier, where the fusion protein induces an anti-IgE immune response  
 CC that does not cause anaphylaxis when administered to an animal. The  
 CC products of the invention have dermatological, antiinflammatory and  
 CC ophthalmological activity. The antigenic peptide described is capable of  
 CC preventing IGE from binding to high affinity receptors on mast cells and  
 CC basophils. The products of the invention are useful in the manufacture of  
 CC a medicament for treating or preventing IGE-mediated allergic disorders  
 CC including asthma, allergic rhinitis, conjunctivitis, glomerular nephritis, flea  
 CC food allergies, eosinophilia, in an animal, e.g. human or dog. The  
 CC polynucleotide products are useful for treating IGE-mediated allergic  
 CC disorders, by gene therapy. Antigenic peptides comprising conserved amino  
 CC acid residues of the CH3 domain of an IGE molecule from one species  
 CC flanked by variable amino acid residues of the CH3 domain of an IGE  
 CC molecule from a second unrelated species are capable of inducing a high  
 CC titre of anti-IgE antibodies when administered to an animal without  
 CC causing anaphylaxis. ABG74772-ABG74785 represents polypeptide sequences  
 CC used in designing the constructs described in the disclosure of the  
 CC invention.  
 XX  
 SQ Sequence 114 AA;  
 Query Match 83.7%; Score 118; DB 6; Length 114;  
 Best Local Similarity 95.7%; Pred. No. 4.7e-10;  
 Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 GETYYSRVTHPLPKDIVRSIAK 24  
 DB 84 GETYYCRVTHPLPKDIVRSIAK 106  
 RESULT 9  
 ABG74774  
 ID ABG74774 standard; protein; 115 AA.  
 XX  
 AC ABG74774;  
 XX  
 DT 23-OCT-2003 (revised)  
 DT 05-JUN-2003 (first entry)  
 XX  
 DE Human IgE CH3 domain-canine CH3 domain chimeric construct.  
 XX  
 CH3 domain; IgE; antigen; non-anaphylactic; anti-IgE; fusion protein;  
 KW dermatological; antiinflammatory; ophthalmological; allergy; asthma;  
 KW allergic rhinitis; gastrointestinal allergy; food allergy; eosinophilia;  
 KW conjunctivitis; glomerular nephritis; flea allergy; atopic dermatitis;  
 KW gene therapy.  
 XX  
 OS Homo sapiens.  
 OS Canis familiaris.  
 OS Chimeric.  
 XX  
 EP1362491-A2.  
 XX  
 04-DEC-2002.

XX 22-MAY-2002; 2002EP-00253606.  
 XX 22-MAY-2001; 2001US-0292638P.  
 XX (PFIZ ) PFIZER PROD INC.  
 XX Brown TM, Morsey MA;  
 PI WPI; 2003-122561/12.  
 DR N-PSDB; ACA55172.  
 XX Novel isolated antigenic peptide comprising amino acid residues of CH3  
 PT domain of IGE molecule from first species and a second unrelated species,  
 PT induces non-anaphylactic anti-IgE immune response in animal.  
 XX  
 PS Claim 2; Page 29-30; 50pp; English.  
 XX  
 CC This invention describes a novel antigenic peptide comprising amino acid  
 CC residues of an IGE CH3 domain from a first species (ADE1) and amino acid  
 CC residues of an IGE CH3 domain of a second unrelated species (ADE2), where  
 CC ADE1 is conserved in the IGE CH3 domain of the second species and ADE2 is  
 CC not conserved in the IGE CH3 domain of the first species. The novel  
 CC antigenic peptide induces a non-anaphylactic anti-IgE immune response in  
 CC an animal. The invention also discloses the polynucleotide sequence  
 CC encoding the antigenic peptide and an antigenic fusion protein comprising  
 CC the antigenic peptide of the invention and a heterologous protein  
 CC carrier, where the fusion protein induces an anti-IgE immune response  
 CC that does not cause anaphylaxis when administered to an animal. The  
 CC products of the invention have dermatological, antiinflammatory and  
 CC ophthalmological activity. The antigenic peptide described is capable of  
 CC preventing IGE from binding to high affinity receptors on mast cells and  
 CC basophils. The products of the invention are useful in the manufacture of  
 CC a medicament for treating or preventing IGE-mediated allergic disorders  
 CC including asthma, allergic rhinitis, conjunctivitis, glomerular nephritis, flea  
 CC food allergies, eosinophilia, in an animal, e.g. human or dog. The  
 CC polynucleotide products are useful for treating IGE-mediated allergic  
 CC disorders, by gene therapy. Antigenic peptides comprising conserved amino  
 CC acid residues of the CH3 domain of an IGE molecule from one species  
 CC flanked by variable amino acid residues of the CH3 domain of an IGE  
 CC molecule from a second unrelated species are capable of inducing a high  
 CC titre of anti-IgE antibodies when administered to an animal without  
 CC causing anaphylaxis. ABG74772-ABG74785 represents polypeptide sequences  
 CC used in designing the constructs described in the disclosure of the  
 CC invention. (Updated on 23-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 115 AA;  
 Query Match 83.7%; Score 118; DB 6; Length 115;  
 Best Local Similarity 95.7%; Pred. No. 4.8e-10;  
 Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 GETYYSRVTHPLPKDIVRSIAK 24  
 DB 85 GETYYCRVTHPLPKDIVRSIAK 107  
 RESULT 10  
 ABG74773  
 ID ABG74773 standard; protein; 117 AA.  
 XX  
 AC ABG74773;  
 XX  
 DT 05-JUN-2003 (first entry)  
 XX  
 DE Human IgE CH3 domain-canine CH3 domain fusion construct.  
 XX  
 CH3 domain; IgE; antigen; non-anaphylactic; anti-IgE; fusion protein;  
 KW dermatological; antiinflammatory; ophthalmological; allergy; asthma;  
 KW allergic rhinitis; gastrointestinal allergy; food allergy; eosinophilia;  
 KW conjunctivitis; glomerular nephritis; flea allergy; atopic dermatitis;  
 KW gene therapy; human; dog.

```

XX OS Unidentified.
XX OS
XX PN WO9967293-Al.
XX PN
XX PD 29-DEC-1999.
XX PF
XX PF 21-JUN-1999; 99WO-US013959.
XX PR
XX PR 20-JUN-1998; 98US-00100287.
XX XX
XX PA (UNBI-) UNITED BIOMEDICAL INC.
XX PI Wang CY, Walfield AM;
XX DR WPI; 2000-160578/14.
XX PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
XX PT for immunization against allergy.
XX PS Claim 14; Page 77; 155pp; English.
XX CC The present invention describes immunoglobulin E (IgE)-CH3 domain
XX CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic and
XX CC anti-asthmatic properties. (I) induces polyclonal antibodies specific for
XX CC a target effector site on the epsilon-heavy chain of IgE, and so
XX CC preventing triggering and activation of mast cells and basophils and
XX CC downregulation of IgE synthesis. Conjugates, or fusion peptides,
XX CC containing (I) are used for active immunisation against IgE-mediated
XX CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
XX CC dermatitis. Nucleic acids that encode these compounds are useful for
XX CC recombinant production of corresponding peptides or in DNA vaccines.
XX CC Conjugates of (I) that include a promiscuous T helper cell epitope
XX CC (functional in genetically diverse subjects), in addition to a B cell
XX CC target epitope, have increased immunogenicity and may include cyclic
XX CC constraints (disulfide bridge) to stabilise conformational features and
XX CC maximize cross-reactivity to the natural target. They induce safe (non-
XX CC anaphylactogenic) antibodies. AAY7994 to AAY80084 represent amino acid
XX CC sequences used in the exemplification of the present invention
XX SQ Sequence 62 AA;

Query Match 100.0%; Score 141; DB 3; Length 62;
Best Local Similarity 100.0%; Pred. No. 7e-14;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
DB 38 CGETYYSRVTHPLPKDIVRSIAKC 62

RESULT 7
AAY80084
ID AAY80084 standard; peptide; 63 AA.
XX
XX AC AAY80084;
XX DT 15-MAY-2000 (first entry)
XX DE
XX DE IgE immunogenic peptide conjugate SEQ ID NO:91.
XX KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
XX KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
XX KW antibody; allergy; allergic disease; immunisation; anti-allergic;
XX KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
XX OS Unidentified.
XX OS
XX PN WO9967293-Al.
XX PD 29-DEC-1999.
XX PF 21-JUN-1999; 99WO-US013959.
XX PF

```

---

```

XX PR 20-JUN-1998; 98US-00100287.
XX PR (UNBI-) UNITED BIOMEDICAL INC.
XX PI Wang CY, Walfield AM;
XX DR WPI; 2000-160578/14.
XX PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
XX PT for immunization against allergy.
XX PS Claim 14; Page 77; 155pp; English.
XX CC The present invention describes immunoglobulin E (IgE)-CH3 domain
XX CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic and
XX CC anti-asthmatic properties. (I) induces polyclonal antibodies specific for
XX CC a target effector site on the epsilon-heavy chain of IgE, and so
XX CC preventing triggering and activation of mast cells and basophils and
XX CC downregulation of IgE synthesis. Conjugates, or fusion peptides,
XX CC containing (I) are used for active immunisation against IgE-mediated
XX CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
XX CC dermatitis. Nucleic acids that encode these compounds are useful for
XX CC recombinant production of corresponding peptides or in DNA vaccines.
XX CC Conjugates of (I) that include a promiscuous T helper cell epitope
XX CC (functional in genetically diverse subjects), in addition to a B cell
XX CC target epitope, have increased immunogenicity and may include cyclic
XX CC constraints (disulfide bridge) to stabilise conformational features and
XX CC maximize cross-reactivity to the natural target. They induce safe (non-
XX CC anaphylactogenic) antibodies. AAY7994 to AAY80084 represent amino acid
XX CC sequences used in the exemplification of the present invention
XX SQ Sequence 63 AA;

Query Match 100.0%; Score 141; DB 3; Length 63;
Best Local Similarity 100.0%; Pred. No. 7.2e-14;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
DB 39 CGETYYSRVTHPLPKDIVRSIAKC 63

RESULT 8
ABG74772
ID ABG74772 standard; protein; 114 AA.
XX
XX AC ABG74772;
XX DT 05-JUN-2003 (first entry)
XX DE
XX DE Canine IgE CH3 domain.
XX KW CH3 domain; IgE; antigen; non-anaphylactic; anti-IgE; fusion protein;
XX KW dermatological; antiinflammatory; ophthalmological; allergy; asthma;
XX KW allergic rhinitis; gastrointestinal allergy; food allergy; eosinophilia;
XX KW conjunctivitis; glomerular nephritis; flea allergy; atopic dermatitis;
XX KW gene therapy; dog.
XX OS Canis familiaris.
XX OS
XX PN EP1262491-A2.
XX PD 04-DEC-2002.
XX PF 22-MAY-2002; 2002EP-00253606.
XX PR 22-MAY-2001; 2001US-0292638P.
XX PA (PFIZ ) PFIZER PROD INC.
XX PI Brown TM, Morsey MA;
XX XX

```

RESULT 4  
 AAY80020  
 ID AAY80020 standard; peptide; 46 AA.  
 AC AAY80020;  
 XX  
 DT 15-MAY-2000 (first entry)  
 XX  
 DE IGE immunogenic peptide conjugate SEQ ID NO:27.  
 XX  
 KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;  
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;  
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;  
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.  
 XX  
 OS Unidentified.  
 XX  
 PN WO9967293-A1.  
 XX  
 PD 29-DEC-1999.  
 XX  
 PF 21-JUN-1999; 99WO-US013959.  
 XX  
 PR 20-JUN-1998; 98US-00100287.  
 XX  
 PA (UNBI-) UNITED BIOMEDICAL INC.  
 XX  
 PI Wang CY, Walfield AM;  
 XX  
 DR WPI; 2000-160578/14.  
 XX  
 PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions  
 PT for immunization against allergy.  
 XX  
 PS Claim 14; Page 77; 155pp; English.  
 XX  
 CC The present invention describes immunoglobulin E (IgE)-CH3 domain  
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic and  
 CC anti-asthmatic properties. (I) induces polyclonal antibodies specific for  
 CC a target effector site on the epsilon-heavy chain of IgE, and so  
 CC preventing triggering and activation of mast cells and basophils and  
 CC downregulation of IgE synthesis. Conjugates, or fusion peptides,  
 CC containing (I) are used for active immunisation against IgE-mediated  
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy  
 CC dermatitis. Nucleic acids that encode these compounds are useful for  
 CC recombinant production of corresponding peptides or in DNA vaccines.  
 CC Conjugates of (I) that include a promiscuous T helper cell epitope  
 CC (functional in genetically diverse subjects), in addition to a B cell  
 CC target epitope, have increased immunogenicity and may include cyclic  
 CC constraints (disulfide bridge) to stabilise conformational features and  
 CC maximize cross-reactivity to the natural target. They induce safe (non-  
 CC anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino acid  
 CC sequences used in the exemplification of the present invention  
 XX  
 SQ Sequence 46 AA;  
 Query Match 100.0%; Score 141; DB 3; Length 46;  
 Best Local Similarity 100.0%; Pred. No. 5e-14; Mismatches 0; Gaps 0;  
 Matches 25; Conservative 0; Indels 0;  
 QY 1 CGETYSRVTHPLPKDIVRSIAKC 25  
 DB 22 CGETYSRVTHPLPKDIVRSIAKC 46  
 RESULT 5  
 AAY80081  
 ID AAY80081 standard; peptide; 57 AA.  
 AC AAY80081;  
 XX

DT 15-MAY-2000 (first entry)  
 XX  
 DE IGE immunogenic peptide conjugate SEQ ID NO:88.  
 XX  
 KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;  
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;  
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;  
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.  
 XX  
 OS Unidentified.  
 XX  
 PN WO9967293-A1.  
 XX  
 PD 29-DEC-1999.  
 XX  
 PF 21-JUN-1999; 99WO-US013959.  
 XX  
 PR 20-JUN-1998; 98US-00100287.  
 XX  
 PA (UNBI-) UNITED BIOMEDICAL INC.  
 XX  
 PI Wang CY, Walfield AM;  
 XX  
 DR WPI; 2000-160578/14.  
 XX  
 PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions  
 PT for immunization against allergy.  
 XX  
 PS Claim 14; Page 77; 155pp; English.  
 XX  
 CC The present invention describes immunoglobulin E (IgE)-CH3 domain  
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic and  
 CC anti-asthmatic properties. (I) induces polyclonal antibodies specific for  
 CC a target effector site on the epsilon-heavy chain of IgE, and so  
 CC preventing triggering and activation of mast cells and basophils and  
 CC downregulation of IgE synthesis. Conjugates, or fusion peptides,  
 CC containing (I) are used for active immunisation against IgE-mediated  
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy  
 CC dermatitis. Nucleic acids that encode these compounds are useful for  
 CC recombinant production of corresponding peptides or in DNA vaccines.  
 CC Conjugates of (I) that include a promiscuous T helper cell epitope  
 CC (functional in genetically diverse subjects), in addition to a B cell  
 CC target epitope, have increased immunogenicity and may include cyclic  
 CC constraints (disulfide bridge) to stabilise conformational features and  
 CC maximize cross-reactivity to the natural target. They induce safe (non-  
 CC anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino acid  
 CC sequences used in the exemplification of the present invention  
 XX  
 SQ Sequence 57 AA;  
 Query Match 100.0%; Score 141; DB 3; Length 57;  
 Best Local Similarity 100.0%; Pred. No. 6.4e-14; Mismatches 0; Gaps 0;  
 Matches 25; Conservative 0; Indels 0;  
 QY 1 CGETYSRVTHPLPKDIVRSIAKC 25  
 DB 33 CGETYSRVTHPLPKDIVRSIAKC 57  
 RESULT 6  
 AAY80080  
 ID AAY80080 standard; peptide; 62 AA.  
 AC AAY80080;  
 XX  
 DT 15-MAY-2000 (first entry)  
 XX  
 DE IGE immunogenic peptide conjugate SEQ ID NO:87.  
 XX  
 KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;  
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;  
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;  
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

CC (functional in genetically diverse subjects), in addition to a B cell  
 CC target epitope, have increased immunogenicity and may include cyclic  
 CC constraints (disulfide bridge) to stabilise conformational features and  
 CC maximize cross-reactivity to the natural target. They induce safe (non-  
 CC anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino acid  
 CC sequences used in the exemplification of the present invention  
 XX  
 XX  
 SQ Sequence 25 AA;

Query Match 100.0%; Score 141; DB 3; Length 25;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-14;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYSRVTHPLPKDIVRSIAKC 25  
 |||||  
 DB 1 CGETYSRVTHPLPKDIVRSIAKC 25  
 |||||

RESULT 2  
 AAY80019  
 ID AAY80019 standard; peptide; 45 AA.  
 AC AAY80019;  
 XX  
 DT 15-MAY-2000 (first entry)  
 XX  
 DE IGE immunogenic peptide conjugate SEQ ID NO:26.  
 XX Immunoglobulin E; IGE; epsilon heavy chain; antigenic; antigen;  
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;  
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;  
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.  
 XX  
 OS Unidentified.  
 XX  
 PN WO9967293-A1.  
 XX  
 PD 29-DEC-1999.  
 XX  
 PF 21-JUN-1999; 99WO-US013959.  
 XX  
 PR 20-JUN-1998; 98US-00100287.  
 XX  
 PA (UNBI-) UNITED BIOMEDICAL INC.  
 XX  
 PI Wang CY, Walfield AM;  
 XX WPI; 2000-160578/14.  
 DR  
 XX  
 PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions  
 PT for immunization against allergy.  
 XX  
 PS Claim 14; Page 76; 155pp; English.  
 XX  
 CC The present invention describes immunoglobulin E (IGE)-CH3 domain  
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic and  
 CC anti-asthmatic properties. (I) induces polyclonal antibodies specific for  
 CC a target effector site on the epsilon-heavy chain of IGE, and so  
 CC preventing triggering and activation of mast cells and basophils and  
 CC downregulation of IGE synthesis. Conjugates, or fusion peptides,  
 CC containing (I) are used for active immunisation against IGE-mediated  
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy  
 CC dermatitis. Nucleic acids that encode these compounds are useful for  
 CC recombinant production of corresponding peptides or in DNA vaccines.  
 CC Conjugates of (I) that include a promiscuous T helper cell epitope  
 CC (functional in genetically diverse subjects), in addition to a B cell  
 CC target epitope, have increased immunogenicity and may include cyclic  
 CC constraints (disulfide bridge) to stabilise conformational features and  
 CC maximize cross-reactivity to the natural target. They induce safe (non-  
 CC anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino acid  
 CC sequences used in the exemplification of the present invention  
 XX  
 XX  
 SQ Sequence 45 AA;

Query Match 100.0%; Score 141; DB 3; Length 45;  
 Best Local Similarity 100.0%; Pred. No. 4.9e-14;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYSRVTHPLPKDIVRSIAKC 25  
 |||||  
 DB 21 CGETYSRVTHPLPKDIVRSIAKC 45  
 |||||

RESULT 3  
 AAY80083  
 ID AAY80083 standard; peptide; 45 AA.  
 AC AAY80083;  
 XX  
 DT 15-MAY-2000 (first entry)  
 XX  
 DE IGE immunogenic peptide conjugate SEQ ID NO:90.  
 XX Immunoglobulin E; IGE; epsilon heavy chain; antigenic; antigen;  
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;  
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;  
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.  
 XX  
 OS Unidentified.  
 XX  
 PN WO9967293-A1.  
 XX  
 PD 29-DEC-1999.  
 XX  
 PF 21-JUN-1999; 99WO-US013959.  
 XX  
 PR 20-JUN-1998; 98US-00100287.  
 XX  
 PA (UNBI-) UNITED BIOMEDICAL INC.  
 XX  
 PI Wang CY, Walfield AM;  
 XX WPI; 2000-160578/14.  
 DR  
 XX  
 PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions  
 PT for immunization against allergy.  
 XX  
 PS Claim 14; Page 77; 155pp; English.  
 XX  
 CC The present invention describes immunoglobulin E (IGE)-CH3 domain  
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic and  
 CC anti-asthmatic properties. (I) induces polyclonal antibodies specific for  
 CC a target effector site on the epsilon-heavy chain of IGE, and so  
 CC preventing triggering and activation of mast cells and basophils and  
 CC downregulation of IGE synthesis. Conjugates, or fusion peptides,  
 CC containing (I) are used for active immunisation against IGE-mediated  
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy  
 CC dermatitis. Nucleic acids that encode these compounds are useful for  
 CC recombinant production of corresponding peptides or in DNA vaccines.  
 CC Conjugates of (I) that include a promiscuous T helper cell epitope  
 CC (functional in genetically diverse subjects), in addition to a B cell  
 CC target epitope, have increased immunogenicity and may include cyclic  
 CC constraints (disulfide bridge) to stabilise conformational features and  
 CC maximize cross-reactivity to the natural target. They induce safe (non-  
 CC anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino acid  
 CC sequences used in the exemplification of the present invention  
 XX  
 SQ Sequence 45 AA;

Query Match 100.0%; Score 141; DB 3; Length 45;  
 Best Local Similarity 100.0%; Pred. No. 4.9e-14;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYSRVTHPLPKDIVRSIAKC 25  
 |||||  
 DB 21 CGETYSRVTHPLPKDIVRSIAKC 45  
 |||||



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 08:16:50 ; Search time 46.6 Seconds  
(without alignments)  
151.581 Million cell updates/sec

Title: US-09-701-623C-6

Perfect score: 141

Sequence: 1 CGETYSRVTHPHLPKDIVRSIAKC 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Genesep 29Jan04:\*

1: Genesep1980s:\*

2: Genesep1990s:\*

3: Genesep2000s:\*

4: Genesep2001s:\*

5: Genesep2002s:\*

6: Genesep2003as:\*

7: Genesep2003bs:\*

8: Genesep2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	141	100.0	25	3	AAY79999 Optimised
2	141	100.0	45	3	AAY80019 IGE immun
3	141	100.0	45	3	AAY80083 IGE immun
4	141	100.0	46	3	AAY80020 IGE immun
5	141	100.0	57	3	AAY80081 IGE immun
6	141	100.0	62	3	AAY80080 IGE immun
7	141	100.0	63	3	AAY80084 IGE immun
8	118	83.7	114	6	ABG74772 Canine Ig
9	118	83.7	115	6	ABG74774 Human IGE
10	118	83.7	117	6	ABG74773 Human IGE
11	118	83.7	124	2	AAW24097 Partial c
12	118	83.7	312	3	AAY79995 Dog immun
13	118	83.7	341	3	AAW06208 Immunogen
14	118	83.7	346	6	ABG74781 Human IGE
15	118	83.7	347	6	ABG74782 Human CH2
16	118	83.7	348	6	ABG74782 Human CH2
17	118	83.7	417	2	AAW23067 Canine Ig
18	118	83.7	426	2	AAW97753 Canine Ig
19	118	83.7	426	6	ABP96583 Dog IGE h
20	108	76.6	25	3	AAY68602 Peptide s
21	108	76.6	25	3	AAY91212 Modified
22	108	76.6	25	3	AAY79998 Optimised
23	108	76.6	42	3	AAY68604 Peptide s
24	108	76.6	42	3	AAY91216 Modified
25	108	76.6	42	3	AAY91215 Modified

26	108	76.6	42	3	AAY91217 Modified
27	108	76.6	42	3	AAY80014 IGE immun
28	108	76.6	44	7	ADD89950 IGE pepti
29	108	76.6	45	3	AAY68605 Peptide s
30	108	76.6	45	3	AAY91218 Modified
31	108	76.6	45	3	AAY80007 IGE-CH3 d
32	108	76.6	45	7	ADD89951 IGE pepti
33	108	76.6	46	3	AAY68603 Peptide s
34	108	76.6	46	3	AAY91213 Modified
35	108	76.6	46	3	AAY91214 Modified
36	108	76.6	46	3	AAY80011 IGE immun
37	108	76.6	56	3	AAY80016 IGE immun
38	108	76.6	59	3	AAY80010 IGE immun
39	108	76.6	60	3	AAY80013 IGE immun
40	108	76.6	60	3	AAY80015 IGE immun
41	108	76.6	63	3	AAY68606 Peptide s
42	108	76.6	63	3	AAY91219 Inv epit
43	108	76.6	63	3	AAY80008 IGE-CH3 d
44	108	76.6	63	3	AAY80012 IGE immun
45	105	74.5	25	3	AAY80077 Optimised

#### ALIGNMENTS

##### RESULT 1

AAY79999 standard; peptide; 25 AA.

XX AAY79999;

AC AC

DT 15-MAY-2000 (first entry)

XX Optimised IGE-CH3 domain antigen peptide for dog IGE.

XX Immunoglobulin E; IGE; epsilon heavy chain; antigenic; antigen;

KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;

KW antibody; allergy; allergic disease; immunisation; anti-allergic;

KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Canis sp.

OS Synthetic.

XX WO9967293-A1.

XX 29-DEC-1999.

PD 21-JUN-1999; 99WO-US013959.

XX 20-JUN-1998; 98US-00100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions

PT for immunization against allergy.

XX Claim 1; Page 99; 155pp; English.

XX The present invention describes immunoglobulin E (IGE)-CH3 domain

CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic and

CC anti-asthmatic properties. (I) induces polyclonal antibodies specific for

CC a target effector site on the epsilon-heavy chain of IGE, and so

CC preventing triggering and activation of mast cells and basophils and

CC downregulation of IGE synthesis. Conjugates, or fusion peptides,

CC containing (I) are used for active immunisation against IGE-mediated

CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy

CC dermatitis. Nucleic acids that encode these compounds are useful for

CC recombinant production of corresponding peptides or in DNA vaccines.

CC Conjugates of (I) that include a promiscuous T helper cell epitope

```
; OTHER INFORMATION: Xaa indicates epsilon-Lys
US-10-355-161A-10

Query Match      71.0%; Score 103; DB 15; Length 44;
Best Local Similarity 68.0%; Pred. No. 2e-07;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGGYQSRVDHPHPKPIVRSITKC 25
Db 20 CGGYQSRVTHPHLPALMRSTKC 44

RESULT 12
US-10-076-674-11
; Sequence 11, Application US/10076674
; Publication No. US20030165478A1
; GENERAL INFORMATION:
; APPLICANT: Sokoll, Kenneth K.
; TITLE OF INVENTION: Stabilized Synthetic Immunogen Delivery System
; FILE REFERENCE: Immunogen Delivery System
; CURRENT APPLICATION NUMBER: US/10/076,674
; CURRENT FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (20)..(20)
; OTHER INFORMATION: Xaa indicates epsilon-Lys
US-10-076-674-11

Query Match      71.0%; Score 103; DB 14; Length 45;
Best Local Similarity 68.0%; Pred. No. 2.1e-07;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGGYQSRVDHPHPKPIVRSITKC 25
Db 21 CGGYQSRVTHPHLPALMRSTKC 45

RESULT 13
US-10-355-161A-11
; Sequence 11, Application US/10355161A
; Publication No. US20040009897A1
; GENERAL INFORMATION:
; APPLICANT: Sokoll, Kenneth K.
; TITLE OF INVENTION: Stabilized Synthetic Immunogen Delivery System
; FILE REFERENCE: Immunogen Delivery System
; CURRENT APPLICATION NUMBER: US/10/355,161A
; CURRENT FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: US 10/076674
; PRIOR FILING DATE: 2002-02-14
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (20)..(20)
; OTHER INFORMATION: Xaa indicates epsilon-Lys
US-10-355-161A-11

Query Match      71.0%; Score 103; DB 15; Length 45;
Best Local Similarity 68.0%; Pred. No. 2.1e-07;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGGYQSRVDHPHPKPIVRSITKC 25
Db 21 CGGYQSRVTHPHLPALMRSTKC 44

RESULT 14
US-09-949-375A-23
; Sequence 23, Application US/09949375A
; Patent No. US20020172673A1
; GENERAL INFORMATION:
; APPLICANT: KLYSNER, Steen et al.
; TITLE OF INVENTION: METHOD FOR DOWN-REGULATING IGE
; FILE REFERENCE: 3631-0111P
; CURRENT APPLICATION NUMBER: US/09/949,375A
; CURRENT FILING DATE: 2002-01-18
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
; LENGTH: 332
; TYPE: PRT
; ORGANISM: mus musculus
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(332)
; OTHER INFORMATION: Murine Ige heavy chain domains C2, C3, and C4.
US-09-949-375A-23

Query Match      67.6%; Score 98; DB 9; Length 332;
Best Local Similarity 82.6%; Pred. No. 8.5e-06;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPIVRSITK 24
Db 191 GYGQCIVDHPDPPKPIVRSITK 213

RESULT 15
US-09-949-375A-25
; Sequence 25, Application US/09949375A
; Patent No. US20020172673A1
; GENERAL INFORMATION:
; APPLICANT: KLYSNER, Steen et al.
; TITLE OF INVENTION: METHOD FOR DOWN-REGULATING IGE
; FILE REFERENCE: 3631-0111P
; CURRENT APPLICATION NUMBER: US/09/949,375A
; CURRENT FILING DATE: 2002-01-18
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial amino acid sequence of SEQ ID NO: 24.
US-09-949-375A-25

Query Match      67.6%; Score 98; DB 9; Length 332;
Best Local Similarity 82.6%; Pred. No. 8.5e-06;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPIVRSITK 24
Db 191 GYGQCIVDHPDPPKPIVRSITK 213

Search completed: February 26, 2004, 08:46:54
Job time : 26.2 secs
```

```
RESULT 7
US-10-176-664-9
; Sequence 9, Application US/10176664
; Publication No. US20030031663A1
; GENERAL INFORMATION:
; APPLICANT: Hellman, Lars T.
; TITLE OF INVENTION: ENHANCED VACCINES
; CURRENT APPLICATION NUMBER: US/10/176,664
; CURRENT FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US/09/401,636
; PRIOR FILING DATE: 1999-09-22
; PRIOR APPLICATION NUMBER: US 60/106,652
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated proteins
US-10-176-664-9
Query Match      84.1%; Score 122; DB 14; Length 341;
Best Local Similarity 95.7%; Pred. No. 3.6e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 GEGYQSRVDHPHPKPIVRSITK 24
      ||||| ||||| ||||| ||||| |||||
Db      206 GEGYQCRVDHPHPKPIVRSITK 228

RESULT 8
US-10-152-190-14
; Sequence 14, Application US/10152190
; Publication No. US20030096369A1
; GENERAL INFORMATION:
; APPLICANT: Morsey, Mohamad A.
; TITLE OF INVENTION: NO. US20030096369A1-anaphylactogenic Ige vaccines
; CURRENT APPLICATION NUMBER: US/10/152,190
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 14
; LENGTH: 346
; TYPE: PRT
; ORGANISM: IgE-5 fusion protein
US-10-152-190-14
Query Match      84.1%; Score 122; DB 14; Length 346;
Best Local Similarity 95.7%; Pred. No. 3.6e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 GEGYQSRVDHPHPKPIVRSITK 24
      ||||| ||||| ||||| ||||| |||||
Db      212 GEGYQCRVDHPHPKPIVRSITK 234

RESULT 9
US-10-214-524-34
; Sequence 34, Application US/10214524
; Publication No. US20030073142A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Swei-Shen Alex
; APPLICANT: Yang, Yong-Min
; APPLICANT: Barankiewicz, Theresa J.
; APPLICANT: Chen, Zhong
; TITLE OF INVENTION: IMMUNOGLOBULIN E VACCINES AND METHODS OF USE THEREOF
; FILE REFERENCE: IGB-00101.P.1.1

; CURRENT APPLICATION NUMBER: US/10/214,524
; CURRENT FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/312,120
; PRIOR FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 34
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Rat (Rattus norvegicus)
US-10-214-524-34
Query Match      84.1%; Score 122; DB 14; Length 428;
Best Local Similarity 95.7%; Pred. No. 4.5e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 GEGYQSRVDHPHPKPIVRSITK 24
      ||||| ||||| ||||| ||||| |||||
Db      283 GEGYQCRVDHPHPKPIVRSITK 305

RESULT 10
US-10-076-674-10
; Sequence 10, Application US/10076674
; Publication No. US20030165478A1
; GENERAL INFORMATION:
; APPLICANT: Sokoll, Kenneth K.
; TITLE OF INVENTION: Stabilized Synthetic Immunogen Delivery System
; FILE REFERENCE: Immunogen Delivery System
; CURRENT APPLICATION NUMBER: US/10/076,674
; CURRENT FILING DATE: 2002-04-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 10
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (19)..(19);
; OTHER INFORMATION: Xaa indicates epsilon-Lys
US-10-076-674-10
Query Match      71.0%; Score 103; DB 14; Length 44;
Best Local Similarity 68.0%; Pred. No. 2e-07;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY      1 CGEGYQSRVDHPHPKPIVRSITKC 25
      ||||| ||||| ||||| ||||| |||||
Db      20 CGETYQSRVTHPLPALMRSTTKC 44

RESULT 11
US-10-355-161A-10
; Sequence 10, Application US/10355161A
; Publication No. US20040009897A1
; GENERAL INFORMATION:
; APPLICANT: Sokoll, Kenneth K.
; TITLE OF INVENTION: Stabilized Synthetic Immunogen Delivery System
; FILE REFERENCE: Immunogen Delivery System
; CURRENT APPLICATION NUMBER: US/10/355,161A
; CURRENT FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: US 10/076674
; PRIOR FILING DATE: 2002-02-14
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 10
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (19)..(19)
```

```
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated proteins
US-09-401-636-2

Query Match      84.1%; Score 122; DB 9; Length 340;
Best Local Similarity 95.7%; Pred. No. 3.6e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 GEGYQSRVDHPHPKPIVRSITK 24
Db      195 GEGYQCRVDHPHPKPIVRSITK 217

RESULT 3
US-10-176-664-2
; Sequence 2, Application US/10176664
; Publication No. US20030031663A1
; GENERAL INFORMATION:
; APPLICANT: Hellman, Lars T.
; TITLE OF INVENTION: ENHANCED VACCINES
; FILE REFERENCE: 10223/006001
; CURRENT APPLICATION NUMBER: US/10/176,664
; CURRENT FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US/09/401,636
; PRIOR FILING DATE: 1999-09-22
; PRIOR APPLICATION NUMBER: US 60/106,652
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated proteins
US-10-176-664-2

Query Match      84.1%; Score 122; DB 14; Length 340;
Best Local Similarity 95.7%; Pred. No. 3.6e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 GEGYQSRVDHPHPKPIVRSITK 24
Db      195 GEGYQCRVDHPHPKPIVRSITK 217

RESULT 4
US-09-401-636-4
; Sequence 4, Application US/09401636
; Patent No. US20010038843A1
; GENERAL INFORMATION:
; APPLICANT: Hellman, Lars T.
; TITLE OF INVENTION: ENHANCED VACCINES
; FILE REFERENCE: 10223/006001
; CURRENT APPLICATION NUMBER: US/09/401,636
; CURRENT FILING DATE: 1999-09-22
; PRIOR APPLICATION NUMBER: US 60/106,652
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated proteins
US-10-176-664-4

Query Match      84.1%; Score 122; DB 14; Length 341;
Best Local Similarity 95.7%; Pred. No. 3.6e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 GEGYQSRVDHPHPKPIVRSITK 24
Db      206 GEGYQCRVDHPHPKPIVRSITK 228

RESULT 5
US-09-401-636-9
; Sequence 9, Application US/09401636
; Patent No. US20010038843A1
; GENERAL INFORMATION:
; APPLICANT: Hellman, Lars T.
; TITLE OF INVENTION: ENHANCED VACCINES
; FILE REFERENCE: 10223/006001
; CURRENT APPLICATION NUMBER: US/09/401,636
; CURRENT FILING DATE: 1999-09-22
; PRIOR APPLICATION NUMBER: US 60/106,652
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated proteins
US-09-401-636-9

Query Match      84.1%; Score 122; DB 9; Length 341;
Best Local Similarity 95.7%; Pred. No. 3.6e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 GEGYQSRVDHPHPKPIVRSITK 24
Db      206 GEGYQCRVDHPHPKPIVRSITK 228

RESULT 6
US-10-176-664-4
; Sequence 4, Application US/10176664
; Publication No. US20030031663A1
; GENERAL INFORMATION:
; APPLICANT: Hellman, Lars T.
; TITLE OF INVENTION: ENHANCED VACCINES
; FILE REFERENCE: 10223/006001
; CURRENT APPLICATION NUMBER: US/10/176,664
; CURRENT FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US/09/401,636
; PRIOR FILING DATE: 1999-09-22
; PRIOR APPLICATION NUMBER: US 60/106,652
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated proteins
US-10-176-664-4

Query Match      84.1%; Score 122; DB 14; Length 341;
Best Local Similarity 95.7%; Pred. No. 3.6e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 GEGYQSRVDHPHPKPIVRSITK 24
Db      206 GEGYQCRVDHPHPKPIVRSITK 228
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 08:27:32 ; Search time 26.2 Seconds  
(without alignments)  
201.482 Million cell updates/sec

Title: US-09-701-623c-7  
Perfect score: 145  
Sequence: 1 CEGYQSRVDHPHFKPIVRSITK 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 21153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000 -  
Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	122	84.1	114	14	US-10-152-190-5
2	122	84.1	340	9	US-09-401-636-2
3	122	84.1	340	14	US-10-176-664-2
4	122	84.1	341	9	US-09-401-636-4
5	122	84.1	341	9	US-09-401-636-9
6	122	84.1	341	14	US-10-176-664-4
7	122	84.1	341	14	US-10-176-664-9
8	122	84.1	346	14	US-10-152-190-14
9	122	84.1	428	14	US-10-214-524-34
10	103	71.0	44	14	US-10-076-674-10
11	103	71.0	44	15	US-10-355-161A-10
12	103	71.0	45	14	US-10-076-674-11
13	103	71.0	45	15	US-10-355-161A-11
14	98	67.6	332	9	US-09-949-375A-23
15	98	67.6	332	9	US-09-949-375A-25

16	98	67.6	332	9	US-09-949-375A-27
17	98	67.6	341	9	US-09-401-636-6
18	98	67.6	341	14	US-10-176-664-6
19	98	67.6	343	9	US-09-949-375A-20
20	98	67.6	343	9	US-09-949-375A-22
21	98	67.6	421	9	US-09-949-375A-28
22	98	67.6	421	14	US-10-214-524-31
23	98	67.6	421	14	US-10-144-188-14
24	98	67.6	432	9	US-09-949-375A-19
25	98	67.6	503	14	US-10-096-840D-2
26	98	67.6	676	14	US-10-096-840D-4
27	83	57.2	426	14	US-10-214-524-26
28	82	56.6	114	14	US-10-152-190-1
29	82	56.6	115	14	US-10-152-190-3
30	82	56.6	117	14	US-10-152-190-2
31	82	56.6	341	9	US-09-401-636-11
32	82	56.6	341	14	US-10-176-664-11
33	82	56.6	346	14	US-10-152-190-10
34	82	56.6	347	14	US-10-152-190-12
35	82	56.6	348	14	US-10-152-190-11
36	82	56.6	426	14	US-10-214-524-28
37	80	55.2	115	14	US-10-152-190-4
38	80	55.2	129	14	US-10-152-190-6
39	80	55.2	222	9	US-09-809-746-2
40	80	55.2	222	10	US-09-809-715-6
41	80	55.2	320	10	US-09-847-208-6
42	80	55.2	323	9	US-09-949-375A-2
43	80	55.2	323	9	US-09-949-375A-4
44	80	55.2	323	9	US-09-949-375A-6
45	80	55.2	330	9	US-09-949-375A-10

## ALIGNMENTS

### RESULT 1

US-10-152-190-5  
; Sequence 5, Application US/10152190  
; Publication No. US20030096369A1  
; GENERAL INFORMATION:  
; APPLICANT: Morsey, Mohamad A.  
; TITLE OF INVENTION: No. US20030096369A1-anaphylactogenic IgE vaccines  
; FILE REFERENCE: PC11011A  
; CURRENT APPLICATION NUMBER: US/10/152,190  
; CURRENT FILING DATE: 2002-05-21  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 114  
; TYPE: PRT  
; ORGANISM: Rat CH3  
US-10-152-190-5

Query Match 84.1%; Score 122; DB 14; Length 114;  
Best Local Similarity 95.7%; Pred. No. 1.1e-09;  
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GEGYQSRVDHPHFKPIVRSITK 24  
Db 84 GEGYQSRVDHPHFKPIVRSITK 106

### RESULT 2

US-09-401-636-2  
; Sequence 2, Application US/09401636  
; Patent No. US2001003843A1  
; GENERAL INFORMATION:  
; APPLICANT: Hellman, Lars T.  
; TITLE OF INVENTION: ENHANCED VACCINES  
; FILE REFERENCE: 10223/006001  
; CURRENT APPLICATION NUMBER: US/09/401,636  
; CURRENT FILING DATE: 1999-09-22  
; PRIOR APPLICATION NUMBER: US 60/106,652